

PhD in Educational Effectiveness

Attainment and progress differences associated with ethnicity, social class and gender throughout the 6th grade of primary schools in Greece and ways to alleviate them: A study of primary school effectiveness

Panayota Korilaki

Thesis submitted for the Degree of Doctor of Philosophy
at the University of London, Institute of Education

2005



Abstract

This study closely follows the school effectiveness paradigm. The design is longitudinal, in that the 6th grade pupils (Greek and foreign/repatriated) were assessed at the beginning and at the end of the school year in language and in mathematics. 54 schools, 27 of which had 2 or more classrooms and 27 of which had only one classroom in the 6th grade (in total 83 classrooms) were sampled.

The findings indicate that foreign/repatriated pupils lagged behind in their attainment compared with their Greek counterparts in both language and mathematics (although the gap was wider for language). Foreign and repatriated pupils made less progress than Greek pupils in mathematics, but no differences in progress were identified for language. Albanian pupils' attainment, and the attainment of other ethnic origin pupils who were grouped into the 'other' ethnic group category, lagged behind the attainment of the majority group. Both Albanian pupils and pupils from 'other' ethnic origin made less progress than Greek pupils in mathematics. Pupils with low socio-economic status had significantly lower attainment than pupils from higher socio-economic status, in both language and mathematics. Pupils with a lower socio-economic status made less progress in mathematics than pupils belonging to a higher socio-economic status. Boys attained significantly lower than girls in both language and mathematics. Boys also made less progress than girls in language.

Significant classroom effects were identified. Classrooms made a difference to pupils' progress rates. The intra-class correlation coefficients or classroom effects in terms of progress were estimated to be 13.5% for mathematics and 13.9% for language. The study found no school effects.

The quantitative part of the study shows that foreign/repatriated pupils underachieved in relation to Greek pupils. Pupils from lower socio-economic background underachieved in relation to pupils from higher socio-economic background and boys underachieved in relation to girls. The finding of underachievement for pupils from foreign and repatriated ethnic background and for pupils from lower social class background can be related to the poor implementation of the current support teaching arrangements provided to foreign/repatriated pupils and to the absence of support schemes catering for the needs of low achieving pupils who in their majority come from a lower social class. Coach classes and

reception classes aim to assist foreign and repatriated pupils with their homework assigned in the main-stream class and they aim to teach them the Greek language. Some recommendations are made about the necessity to extend the provision of support teaching.

Table of Contents

PhD in Educational Effectiveness	1
Attainment and progress differences associated with ethnicity, social class and gender throughout the 6 th grade of primary schools in Greece and ways to alleviate them: A study of primary school effectivenessl	
Abstract	2
Table of Contents	4
List of tables	7
List of figures	8
List of graphs	9
Appendixes (in a second volume)	10
Glossary	11
Acknowledgements:	12
Introduction	13
1. LITERATURE REVIEW	16
A) THE SCHOOL EFFECTIVENESS RESEARCH TRADITION	16
1. 1. Main uses of School Effectiveness Research	16
1. 2. Orientations within School Effectiveness Research	18
1. 2. 1. The 'equality of opportunities' research tradition	19
1. 2. 2. Describing the remaining four orientations within school effectiveness research	34
1. 3. School and classroom effects	37
1. 4. Differential effectiveness for different student groups	41
1. 5. Identifying links between School Effectiveness and School Improvement	45
1. 6. Transferability of constructs and variables	49
1. 7. Case studies based on a SER design	50
1. 8. Other issues:	51
1. 8.1. Size and significance of effects	52
1. 8. 2. Consistency across outcomes	52
1. 8. 3. Stability over time	54
1. 8. 4. Long term effects of schools	57
1. 9. Outcomes measured	59
1. 10. Other School Effectiveness Criticisms	60
B) UNDERACHIEVEMENT OF ETHNIC MINORITY PUPILS AND PUPILS FROM LOWER SOCIAL CLASS AND EFFECTIVE SUPPORT SCHEMES AIMING TO ALLEVIATE IT	69
1. 11. Describing underachievement of ethnic minority pupils and pupils from lower social class	69
1. 12. Scope of support education	77
1. 13. A summary of support teaching provision in Greece	86
1. 14. Support education programmes in other countries	91
1. 14. 1. Title I	91
1. 14. 2. Success for All	97
1. 14. 3. Reading Recovery	101
1. 14. 4. Levine's approach	104
1. 14. 5. Programmes supporting Numeracy	107
1. 15. Summary and relevance of the literature review for the current study	108
2. METHODOLOGY	111
2. 1. Aims of the study	111
2. 2. Research Design	113
2. 3. Research questions	115
2. 4. Analysis techniques employed in the quantitative part	118
2. 5. Describing pupils' characteristics	125
2. 6. Structure and properties of the performance-based tests employed in the current study and threats to validity associated with the test construction	143
2. 7. Other factors that may reduce the validity of the current study	146
2. 8. Sample	146

2. 9. Response rates	150
2. 10. Numbers of foreign and repatriated pupils	151
2. 11. Confidentiality issues	152
2. 12. The need to employ disaggregated (pupil level) data	153
2. 13. Testing for school or classroom effects	155
2. 14. Contextual characteristics of the classrooms	156
2. 15. Duration of the study	157
2. 16. Time – frame	157
2. 17. Training of research assistants	157
2. 18. Choice of Region	158
2. 19. Choice of the 6th grade	159
2. 20. Pilot study	160
2. 21. Methodological Summary	162
3. MULTILEVEL ANALYSIS OF PRIMARY SCHOOL EFFECTIVENESS IN MATHEMATICS AND IN LANGUAGE	163
3. 1. Issues that will be covered in this analysis	163
3. 2. Multi – level modelling	164
3. 3. Examining the distributions of mathematics tests	166
3. 4. Examining the distributions of language tests	168
3. 4. Examining the distributions of language tests	169
3. 5. Descriptive Statistics of all variables employed in the analysis	170
3. 6. Types of Models Employed in the Analysis	175
3. 7. Null models	181
3. 8. Evaluating the relative standing of classrooms according to progress models with random intercepts	191
3. 9. Drawing graphs of predicted classroom lines according to progress models with random intercepts	195
3. 10. Evaluating the relative standing of classrooms according to a progress model with random coefficients (progress model 2) for mathematics	197
3. 11. Variance function for mathematics at the classroom level according to progress model 2	202
3. 12. Variance at the pupil level for mathematics	204
3. 13. Variance Partition Coefficient	204
3. 14. Evaluating the relative standing of classrooms according to a progress model with random coefficients (progress model 4) for language	207
3. 15. Variance function for language at the classroom level	210
3. 16. Variance function for language at the pupil level	211
3. 17. Variance Partition Coefficient for language	213
3. 18. Predictions according to progress model 2 for mathematics	215
3. 19. Predictions according to progress model 4 for language	221
3. 20. Predictions according to progress model 2 for mathematics and	224
progress model 4 for language	224
3. 21. Investigating the ‘school effect’	227
3. 22. Comparing classroom raw and normalised final performance levels	233
3. 23. Classroom residuals	235
3. 24. Classroom effects	237
3. 25. Examining the classroom residuals derived from null and progress models	241
3. 26. Comparing classroom residuals derived from null models	243
3. 27. Comparing classroom residuals derived from progress models	245
3. 28. Information that classroom residuals can provide	251
3. 29. Checking model assumptions	254
3. 29. 1. Checking the assumption of normality for level 2 residuals	254
3. 29. 2. Checking the assumption of constant variance among level 2 residuals	257
3. 29. 3. Checking the assumption of normality for level 1 residuals	259
3. 29. 4. Checking the assumption of constant variance for level 1 residuals	261
3. 29. 5. Checking the linearity assumption	262
3. 30. Profiling classrooms	263

3. 31. Is it legitimate to make comparisons between classrooms?	270
3. 32. Social class and Mother's education	272
3. 33. Differential effectiveness.....	275
3. 34. Models accounting for individual pupils' characteristics	276
3. 35. Predictions according to the above 'individual characteristics model 3d2'	279
3. 36. Contextual Models	282
3. 37. Are classrooms consistently effective according to a range of models?.....	291
3. 38. Consistency between outcomes	298
3. 39. School Effectiveness Research and concerns of Accountability	303
4. FINDINGS	311
4. 1. The existence of classroom effects.....	311
4. 2. Discussion on the absence of school effects for the current Greek dataset	312
4. 3. Estimating the impact of individual pupils' characteristics on pupils' final attainment scores and progress rates.....	319
4. 3. 1. Differences in attainment and in progress between foreign/repatriated and majority pupils....	329
4. 3. 2. The Joint Effect of Ethnicity and Social Class	341
4. 3. 3. The impact of social class on pupils' attainment and progress.....	345
4. 3. 4. Attainment and progress differences between boys and girls.....	351
4. 3. 5. The Impact of Mothers' Educational Level.....	355
4. 3. 6. The Impact of friends	357
4. 3. 7. The Impact of Nursery School	358
4. 3. 8. The Impact of Homework	360
4. 3. 9. Interaction terms in absolute attainment models	362
4. 3. 10. Interaction terms in progress models.....	365
4. 3. 11. Random coefficients in absolute attainment models	367
4. 3. 12. Random coefficients in progress models.....	370
4. 4. Contextual Effects	371
4. 5. The implementation of support schemes	381
4. 6. The importance of the provision of effective learning time	385
5. CONCLUSIONS.....	389
5. 1. The absence of school effect	389
5. 2. Conditions associated with educational disadvantage	390
5. 2. a. Implications stemming from the underachievement of foreign/repatriated pupils	392
5. 2. b. Implications stemming from the underachievement of pupils from lower social class.....	393
5. 3. The gender issue.....	393
5. 4. Ways to counteract educational disadvantage	394
5. 5. The impact of other variables	401
5. 6. Significant interactions in attainment and progress models	401
5. 7. Implications for further study.....	403
REFERENCES.....	410

List of tables

Table 1. 1: Variance Partition Coefficients for classrooms and schools for different countries from Scheerens et al. (1989)	38
Table 1. 2: Results of ρ and school-level variances for parallel classes in primary schools	40
Table 2. 1: Pupils' citizenship.....	127
Table 2. 2: Detailed Goldthorpe social class schema	138
Table 2. 3: Minimum, maximum and mean number of foreign and repatriated pupils in the schools that took part in the study	151
Table 2. 4: Role, purposes, instruments involved and analyses undertaken in the pilot and the main study	161
Table 3. 1: Mean and standard deviations of variables measuring individual pupils' characteristics ...	170
Table 3. 2: Representation of Greek, Albanian, pupils coming from the democracies of the former USSR, as well as from 'Other' ethnic backgrounds, grouped together in the same category, assessed in mathematics at the end of the school-year.	171
Table 3. 3: Frequencies of pupils' social class categories according to the 9-category Goldthorpe social class schema. Each pupil is allocated to the higher social class between the two social class positions of his/her parents	171
Table 3. 4: Gender by dichotomous social class for Greek pupils: case processing summary.....	172
Table 3. 5: Cross-tabulation of gender by dichotomous social class.....	172
Table 3. 6: Cross-tabulation of ethnic group membership by gender.....	173
Table 3. 7: Cross-tabulation of Ethnic group by dichotomous social class (1: manual, 0: non-manual).....	173
Table 3. 8: Pearson's Correlations between pairs of variables employed in the analysis	174
Table 3. 9: Predictions for low, average and high achievers according to progress model 2 for two classrooms belonging to the same school.....	218
Table 3. 10: Predictions for low achieving, average achieving and high achieving pupils for two classrooms belonging to the same school. All predictions are based on progress model 2.....	220
Table 3. 11: Predictions for low, average and high achievers according to progress model 4 for two classrooms belonging to the same school.....	223
Table 3. 12: Range in classroom mean raw scores at the end of the 6 th grade.	233
Table 3. 13: Range in classroom mean normalized scores at the end of the 6 th grade.	234
Table 3. 14: Fixed part coefficients of social class derived from absolute attainment model 11 for mathematics according to the 9-category Goldthorpe schema	272
Table 3. 15: Classification scheme for mother's educational level	274
Table 3. 15: Classification scheme for mother's educational level	274
Table 3. 16: Descriptive statistics concerning compositional and contextual characteristics of classrooms	283
Table 3. 17: Correlations of intercept residuals derived from the null, the progress, and the progress model further adjusting for individual pupils' characteristics and the contextual model for mathematics.	293
Table 3. 18: Correlation coefficients of intercept residuals derived from the null, the progress, and the progress model adjusting for individual pupils' characteristics and the contextual model for language.....	295
Table 3. 19: Correlation coefficients between mathematics and language intercept residuals derived from null, progress, progress models adjusted for individual pupils' characteristics and contextual models.....	299
Table 4. 1. Adjusted attainment models: The main effects of variables measuring individual pupils' characteristics (in the diagonal); and the percentage reduction of the main effects (in the rows, off the diagonal) when the model controls for a second explanatory variable.....	323
Table 4. 2. Adjusted attainment models: The main effects of variables measuring individual pupils' characteristics (in the diagonal); and the percentage of reduction of the main effects (in the rows, off the diagonal) when the model controls for a second explanatory variable.....	323
Table 4. 3. Adjusted progress models: The main effects of variables measuring individual pupils' characteristics (in the diagonal); and the percentage of reduction of the main effects (in the rows, off the diagonal) when the model controls for a second explanatory variable.....	324

Table 4. 4. Adjusted progress models: The main effects of variables measuring individual pupils' characteristics (in the diagonal); and the percentage of reduction of the main effects (in the rows, off the diagonal) when the model controls for a second explanatory variable.....	324
Table 4. 5. Examining the impact of gender, social class and ethnicity on pupils' attainment and progress rates on both subjects.	326
Table 4. 6. Examining the impact of belonging to a foreign/repatriated ethnic group on pupils' final attainment scores and progress rates for both subjects.....	334
Table 4.7. Examining the impact of social class on pupils' final attainment scores for mathematics and language (the detailed Goldthorpe social class scheme is employed).	348
Table 4. 8. Examining interactions in absolute attainment models	363
Table 4. 9: Significant interaction effects in progress models.....	365
Table 4. 10. Random effects in absolute attainment models	368
Table 4. 11. Contextual models.....	377

List of figures

Figure 1: Null model 2 for mathematics.....	182
Figure 2: Null model 2 for language	184
Figure 1: Null model 2 for mathematics.....	187
Figure 3: Null model 1 for mathematics.....	187
Figure 2: Null model 2 for language	188
Figure 4: Null model 1 for language	188
Figure 5: Progress model 1 for mathematics.....	192
Figure 6: Progress model 1 for language.....	192
Figure 7: Progress model 2 for mathematics.....	200
Figure 8: Progress model 4 for language.....	209
Figure 9: Progress model 3d2 adjusting for pupils' initial score and for pupils' social class, ethnicity (being of foreign ethnic background), and for the interaction term 'being of a foreign/repatriated ethnic origin by initial attainment score'	277
Figure 10: Contextual model cont3 for language, adjusting for the percentage of pupils who are underachieving in the classroom along with initial attainment score, social class, and interaction term 'social class by initial attainment score'	288

List of graphs

Graph 1: End score predictions according to ‘progress’ model 1, for mathematics.....	196
Graph 2: End score predictions according to progress model 4 for language	196
Graph 3: Level-2 variance function by initial attainment score according to progress model 2 for mathematics.....	203
Graph 4: Variance Partition Coefficient by initial attainment score for mathematics, according to progress model 2.....	206
Graph 5: Level-2 variance function by initial attainment score for language according to progress model 4.....	210
Graph 6: Variance function at the pupil level (level 1) by initial attainment score for language according to progress model 4.....	212
Graph 7: Variance Partition Coefficient by initial attainment score for language according to progress model 4.....	213
Graph 8: Final attainment score predictions for each classroom according to ‘progress’ model 2 for mathematics.....	224
Graph 9: Final attainment score predictions according to ‘progress model 4’ for language	225
Graph 10: Classrooms ranked according to null model 2 in mathematics (including classroom and pupil levels). 95% confidence intervals are applied.	243
Graph 11: Classrooms ranked according to null model 2 for language (including classroom and pupil levels). 95% confidence intervals are applied.	243
Graph 12: Classrooms ranked according to their intercept and slope residuals derived from progress model 2 in mathematics. 95% confidence intervals are applied. Intercept residuals enable comparisons between classrooms in terms of progress whereas slope residuals enable comparisons between classrooms in terms of degree of differentiation.	247
Graph 13: Classrooms ranked according to the intercept residuals and to their slope residuals of the progress model 4 for language. 95% confidence intervals are applied.	249
Graph 14: Standardised residuals by normal scores for classroom intercept and slope residuals for mathematics.....	255
Graph 15: Level 2 standardised intercept and slope residuals and their normal scores for language....	256
Graph 16: Standardised residuals by fixed-part predictions at level 2 at average initial attainment score for mathematics.	257
Graph 17: Standardised intercept and slope residuals at level 2 by fixed-part prediction for the average initial score of each class for language.....	258
Graph 18: Standardised residuals by normal scores at level 1 derived from progress model 2 for mathematics.....	259
Graph 19: Standardised residuals by normal scores at level 1 for ‘progress model 4’ for language.....	260
Graph 20: Level 1 standardised residuals by fixed part prediction for mathematics.....	261
Graph 21: Level 1 standardised residuals by fixed part prediction for language	261
Graph 22: Scatter plot of classroom intercept residuals according to progress model 2 (showing progress) with intercept residuals according to null model 2 (showing absolute attainment) for mathematics.	265
Graph 23: Scatter plot of classroom intercept by slope residuals according to progress model 2 for mathematics.....	268
Graph 24: Scatter plot of classroom intercept by slope residuals according to progress model 4 for language.	269
Graph 25: Variance Partition Coefficient by initial attainment score for mathematics according to progress model 3d2 that accounts for individual pupils’ characteristics.....	278
Graph 26: Average predicted lines according to progress model 3d2 adjusting for individual pupils’ characteristics for different categories of pupils for mathematics.....	281

Appendixes (in a second volume)

APPENDIXES	1
A) Table of Contents	2
B) List of tables	11
APPENDIX 1	12
CREATING THE PERFORMANCE-BASED TESTS ON LANGUAGE AND MATHEMATICS	13
A) PURPOSE OF THE TESTS.....	13
B) ADMINISTRATIVE REQUIREMENTS:	18
C) SAMPLE SIZE, LOCATION AND RESPONSE RATES.....	19
D) DESIGNING THE INSTRUMENTS	21
D.1. RATIONALE OF MATHEMATICS TESTS	21
D.1.2. Objectives of mathematics tests.....	23
D.1.3. Characteristics of the distributions of mathematics tests	29
a) Total score distributions in mathematics for all pupils.....	30
b) Total score distributions in mathematics for foreign and repatriated pupils	31
D.2. RATIONALE OF LANGUAGE TESTS	32
D.2.1. CONSTRUCTING LANGUAGE TESTS.....	50
D.2.2. Characteristics of the distributions of language tests.....	54
D.3. RELIABILITY	60
D.3.1. RELIABILITY COEFFICIENTS FOR PERFORMANCE-BASED TESTS ON MATHEMATICS AND LANGUAGE	62
D.3.1.1. Factors affecting reliability.....	68
D.4. STANDARD ERROR OF MEASUREMENT FOR PERFORMANCE-BASED TESTS ON MATHEMATICS AND LANGUAGE	69
D.4.1. STANDARD ERROR OF MEASUREMENT OF MATHEMATICS TESTS	71
D.4.2. STANDARD ERROR OF MEASUREMENT OF LANGUAGE TESTS	75
D.5. ITEM DIFFICULTY	80
D.6. ITEM DISCRIMINATION	81
D.7. ITEM ANALYSIS.....	83
D.8. TEST EQUATING	85
D.9. NORMS.....	86
D.10. BLUEPRINTS – EQUATING	103
2 nd APPENDIX.....	105
Rankings of ‘most effective’ and ‘least effective’ classrooms in a) mathematics and b) language according to a range of models	105
3 rd APPENDIX	118
Principal’s Questionnaire	119
Support Class Teacher’s Questionnaire.....	123
Questionnaire for foreign / repatriated pupils.....	127
Performance-based tests employed to assess all pupils in a) mathematics and b) language.....	121

Glossary

Coach classes provide support to foreign and repatriated pupils with limited competence in Greek in the afternoon, after the end of the school day. They provide additional support to foreign/repatriated pupils who are in the initial stages of learning Greek, so that these pupils become more able to comprehend mainstream class lessons. Some support class teachers in coach classes may instead teach foreign/repatriated pupils Greek as a second language and they provide assistance to the mainstream curriculum only sporadically.

The aim of ***reception classes*** is to assist foreign/repatriated pupils to understand the lessons taught in the mainstream class and to teach them the Greek language. Unlike coach classes, reception classes operate as pull - out classes in the mornings, or in times parallel with the mainstream class. Therefore pupils enrolled in them may miss out important lessons taught in the mainstream. Greek law N1404/83 no 45 institutionalised reception classes and coach classes.

Whole day school: The whole day school operates during after school hours in the afternoon. In its framework '*sections with extended timetable*' are established, in which pupils are *supported with their homework and/ or extra-curricular activities are organised*'. *Initially the whole day school aimed to meet the needs of pupils with working parents, but in most cases sections with extended timetable address the learning needs of all pupils enrolled in the school.* Whole day schools and sections with extended timetable are established by the legislation 2525/97 in ΦΕΚ 188/2Α' article 4.

Acknowledgements:

I would like to express considerable thanks to my main supervisor professor Harvey Goldstein for his assistance in carrying out this thesis, which is a School Effectiveness Project. Also I would like to thank professor Peter Blatchford for assisting me in carrying out the qualitative part of the project as well as professor Jagdish Singh Gundara for having me as his student during the early years of this degree.

I would like to thank the University of London and more specifically the Central Research Fund, as it funded my research project with the sum of 1000 pounds. This included mainly printing expenses and mainly covered the cost of printing the performance-based tests.

On the Greek side, I would like to mention the contribution of the Greek State Scholarship Foundation in providing me with the scholarship, and the substantial help in terms of paying the academic fees as well as the living expenses in the UK for three years and a half. I would like also to mention the contribution of my service (the National Statistical Service of Greece) in financing me for five years, and enabling me to carry out my master dissertation and thesis.

I am also grateful to Professor Lemonidis at the University of Thessaloniki who assisted in the revision of the performance-based tests in mathematics, after the initial piloting phase.

Introduction

The current study deals with primary school and classroom effectiveness. It examines pupils' performance in mathematics and in language over one school year. 1858 Greek and foreign/repatriated pupils enrolled in 83 6th-grade classrooms situated in 54 primary schools were examined twice. The final attainment and the progress rates of foreign/repatriated pupils in relation to the majority group are examined.

The more school years foreign/repatriated pupils are exposed to the Greek language medium and the more years they are taught the Greek language, the more they become competent in the Greek language. Therefore differences in attainment between foreign/repatriated and Greek pupils are expected to decrease the more school years the former spend in the Greek school. Absence of such differences in attainment and in progress would imply that foreign/repatriated pupils have successfully been integrated into the Greek school. Such a situation would imply that foreign/repatriated pupils have equal access to the curriculum taught in the mainstream class, and can comprehend what is being taught, since their performance is similar to that of their Greek counterparts. Attainment and progress differences between pupils from disadvantaged and non-disadvantaged socio-economic backgrounds also constitute cause for concern.

School effectiveness research methodology offers a well-known framework, according to which differences in attainment and progress between schools and/or between classrooms can be identified; similarly significant attainment and/or progress differences between discrete pupil groups can be estimated. On that premise, subsequent remedial action can be undertaken in order to bridge the gap between the educationally disadvan-

tagged and the advantaged. These actions may be perceived within a comprehensive framework of support teaching policies, initiated in order to tackle both these structural inequalities through a systematic and sustained endeavour to alleviate educational and social disadvantage.

Multilevel modelling was chosen as the major analytical tool as it separates out the variation that can be attributed to pupils' characteristics, from the variation that can be attributed to school and classroom membership. It can map out the impact of a range of different variables such as characteristics of individual pupils and characteristics of classrooms and schools on pupils' attainment and on their progress. Characteristics of the pupils enrolled in a given classroom, along with the characteristics of a given classroom and the characteristics of a given school, jointly shape the performance of foreign/repatriated pupils in particular and of disadvantaged groups in general, as well as the performance of majority pupils.

Another issue examined in the current analysis is whether school and / or classroom effects exist, or to what extent schools or classrooms make a difference to pupils' attainment or to their progress rates. This analysis is necessary because pupils' final attainment scores are jointly defined by their individual characteristics and by their school or classroom membership. Furthermore the magnitude of these effects will be estimated reflected in the percentage of the total variation in pupils' final attainment scores that can be attributed to the fact that pupils are enrolled in schools or classrooms.

This study includes a review of literature on School Effectiveness Research and on School Improvement in order to draw lessons from previous research. The chapter on

Methodology includes details about the sample. The quantitative analysis chapter explains in detail some of the multilevel modelling procedures undertaken and it underlines discrepancies in educational attainment and in progress rates between different pupil groups and to show some variables that are influential in shaping pupils' educational outcomes. The findings reveal different patterns of underachievement for foreign and repatriated pupils, for pupils from lower social class background and for boys. Then the issue of implementation of support schemes for foreign and repatriated pupils is described, according to the accounts of principals and support class teachers in the schools investigated.

1. LITERATURE REVIEW

The current literature review is divided into two parts. The first part describes the school effectiveness research tradition and looks more thoroughly into the equal educational opportunities strand. The second part looks into the issue of the underachievement of ethnic minority pupils from a broad perspective and describes effective support schemes that have been adopted internationally in order to alleviate educational and social disadvantage.

A) THE SCHOOL EFFECTIVENESS RESEARCH TRADITION

1. 1. Main uses of School Effectiveness Research

According to Scheerens and Bosker (1997, p. 1) effectiveness may be defined in terms of various criteria and a range of elements and aspects of schooling may be chosen to point at the impact of action aimed at increased performance.

By measuring the performance of individual pupils, the average performance of classrooms and schools participating in School Effectiveness Research may be estimated. Performance can be expressed as a snapshot at a specific point in time (a cross-sectional design) or over time (usually one school year), in a longitudinal design. The performance of classrooms and schools both in terms of attainment and in terms of progress may be further adjusted for individual pupils' characteristics related to conditions of disadvantage. This exercise is undertaken in order to show the *relative standing of the classroom or school* in relation to classrooms or schools in the same area in specific

subjects, and in terms of attainment and progress at the district level, at the regional, or at the national level. In addition, assessment can be the basis for educational monitoring so that *educational inequalities* are alleviated through all agents' joint efforts. Murphy (1992, cited in Kavouri, 1996, p. 99) pointed out that the '*quality of education in terms of organisational, instructional changes and improvement efforts could be assessed by examining student outcomes*'. According to the Organisation for Economic Cooperation and Development (OECD, 1955) cited in Verdis (2002, p. 45) pupils' outcomes can be considered an indicator associated with the *quality of schools* and the *quality of the educational system*. The other four indicators are: the flow of students through the education system, the schools and their environment, the costs of education and students' attitudes and expectations.

The *criteria for an adequate study of school effectiveness* have been described by Scheerens (1992). According to Scheerens, a school effectiveness study:

- '*Taps sufficient 'natural' variance in school and instructional characteristics, so that there is a fair chance that they might be shown to explain differences in achievement between schools;*
- *Uses adequate operationalizations and measures of the process and effect variables, and a mixture of quantitative and qualitative measures;*
- *Adequately adjusts effect measures for intake differences between schools (e.g. in previous achievement and socio-economic status of students);*

- *Has units of analysis that allow for data analysis - in many cases multilevel models will be appropriate to do justice to the fact that we usually look at classes within schools, students within classes and perhaps even schools within specific types of environments;*
- *Uses longitudinal data'.*

School effectiveness research has associated effective classrooms and/or effective schools with practices and policies at the classroom level or at the school level, known as 'effective school correlates'.

1. 2. Orientations within School Effectiveness Research

The following orientations or research traditions in the field of School Effectiveness research have been identified:

The equality of opportunities research tradition, which has been initiated by the Coleman report (Coleman et al., 1966), economic studies on education production functions, the evaluation of compensatory programmes, studies on effective schools and the evaluation of school improvement programmes, and studies on the effectiveness of teachers, classes and instructional processes (Scheerens, 1997, p. 139).

1. 2. 1. The ‘equality of opportunities’ research tradition

In 1966 in U.S.A. Coleman’s famous ‘Equality of Educational Opportunity’ study, better known as the Coleman Report (Coleman et al., 1966) claimed that schools did not make a difference to students’ achievement. Coleman discovered that the impact of ethnicity was interrelated with the school’s impact in shaping students’ outcomes. Coleman found out that school effects for black students were almost twice as large as school effects for white students. Coleman’s study underlined the impact of ethnicity or of social class on pupils’ educational attainment. According to Sheerens and Bosker (1997, p. 143) this study can be perceived as the first of the tradition within School Effectiveness studies focusing on unequal educational opportunities, that is to say the limitations imposed by the environment on pupils’ learning outcomes. The report spurred researchers to try to do a better job of gathering data, asking questions and presenting results. Coleman concluded that there were significant differences in achievement associated with ethnicity, between the American Blacks and the Whites. These differences were also associated with family background and the area where the school was located (OECD, 1980, p. 19). In other words, effects associated with ethnicity, social class and the school’s location were put forward to account for differences in educational attainment. The Coleman study went a step further than the previous sociologically oriented research carried out in 1960s, according to which attainment differences could be attributed only to the differences of pupils’ socio-economic conditions or to differences in their cultural background. According to Scheerens and Bosker (1997) other studies highlighting unequal opportunity in the USA were those of Jenks et al.

(1979), who found similar results to Coleman's study (*op. cit.*), of Alexander & Eckland (1980) and of Hauser et al. (1976).

Measures of low income (e.g. unemployment, eligibility for free school meals, receipt of clothing grants, etc.), large family size one-parent status, and poor housing conditions have also been found to be powerful predictors of academic achievement (e.g. Douglas, 1964; Davie et al., 1972; Rutter & Madge, 1976; Essen & Wedge, 1982; Mortimore and Blackstone, 1982).

School effectiveness research started in the early 1980's in the U.K., in the Netherlands, in U.S.A. etc. The first School Effectiveness studies were quite optimistic about identifying effective school correlates. In the USA the first studies carried out through the school effectiveness paradigm were Brookover et al. (1979), and Edmonds (1979) and in the UK Rutter et al. (1979). Edmonds commented on the findings of his own research (Edmonds, 1978, 1979a, 1979b) and that of others (e.g. Lezotte and Bancroft, 1985; Weber, 1971) to argue for the creation of effective schools for the urban poor.

Edmonds (1982) as well as the majority of School Effectiveness studies emphasised mastery of *basic skills* (pupils' performance in language and mathematics). '*Basic skills provide the foundation for further learning and therefore can be seen as lying at the core of subsequent educational outcomes and progress*' (Kaluge, 1998, p. 161). Relatively few more recent School Effectiveness studies examine the impact of school on pupils' social and affective outcomes (e.g. Brandsma and Knuver, 1989; Cuttance, 1987; Gray, McPherson and Raffe, 1983 cited in Sammons, 1996).

Brookover et al.'s (1979) study, after identifying outlier¹ schools, attempted to explain how more and less effective schools varied in school routines and classroom practices.

Brookover (*op. cit.*) concluded that the combination of schools' racial structure variables accounted for more than 85 per cent of the between school variance in mean reading and mathematics achievement. This study is notable for the inclusion of process variables that could measure school 'climate'.

However, a little later on, Goodlad (1984) proposed a variety of different goals for education. These broader goals included *academic development, intellectual development, vocational skills, social, civic and cultural goals and personal goals*.

In the USA Firestone (1991) reviewed early School Effectiveness studies and highlighted the point that the effective schools movement was committed to the belief that '*children of the urban poor could succeed in school*' (e.g. Edmonds, 1979; Goodlad et al., 1979). Other studies in the 1980s also identified significant differences between ethnic groups in educational attainment at primary school (Mabey, 1981, 1986; Scarr et al., 1983; Osborn and Butler, 1985; Maughan & Dunn, 1988; Mackintosh et al., 1988; Kysel, 1988 cited in Strand, 1999). In Britain, a major longitudinal study was the Junior School Project (JSP) (Mortimore et al., 1988), which looked at 2000 7-year-old pupils for three years in 50 London schools. The JSP followed the progress in reading and mathematics of junior school pupils between the ages of seven and ten years.

¹ Outlier schools are schools having performed at an above average or at a below average level in terms of pupils' progress or in terms of attainment or in terms of adjusted progress or adjusted attainment.

The study also collected data on many individual pupils' characteristics such as pupils' ethnicity and social class, parents' occupation, eligibility for free school meals (low family income), family size, nursery school experience and incomplete fluency in English, in order to allow all these factors' impact on pupils' final performance and progress rates to be assessed simultaneously. Mortimore et al. (1988, p. 186) reported that for reading progress while '*in particular schools, children made greater progress in reading than expected, given their initial attainment and background characteristics, in other schools, children made less progress than expected*'. However, while effective schools promoted the achievement levels of all pupils, they did not narrow the gap between the achievement levels of the disadvantaged and the non-disadvantaged groups. In this study, data were not aggregated at the school's level but students' cases were introduced for the analysis.

The JSP (*op. cit.*) detected differences in reading, writing and mathematics attainment associated with social class. It also detected differences in progress for pupils from lower social class background in reading and writing but no progress differences in mathematics. Furthermore, the JSP identified significant differences in reading and writing attainment between boys and girls, and in favour of girls, but sex did not have a significant effect on progress, either in relation to reading, or in relation to writing. In addition, significant differences between boys and girls for mathematics in terms of attainment and in terms of progress were detected, in favour of girls. Pupils from foreign ethnic backgrounds with inadequate fluency in English performed less well and made less progress than English, Scottish, Welsh and Irish pupils. But again, different attainment and progress patterns were identified between different ethnic groups. For

mathematics, pupils from certain ethnic groups had lower attainment than English, Scottish, Welsh and Irish pupils. Nevertheless, there was no evidence that belonging to a foreign/repatriated ethnic group would have a hindering effect on pupils' progress for mathematics. A follow-up study of the JSP sample to age 16 (Sammons, 1995) identified also attainment differences associated with gender and ethnic groups in primary school but indicated that the gap in attainment associated with economic disadvantage and sex widened even further over the secondary school years. The later study examined the size and stability of gender, and ethnic and socio-economic differences in pupils' educational achievement over a 9-year period. Towards the end of Junior Education background effects associated with individual pupils' characteristics remained fairly stable over the junior age period and they accounted for 20% of the total variance for reading and 11.3% of the total variance for mathematics attainment at the end of age 10. These results showed that the impact of home influences on mathematics attainment is less pronounced than it is on reading attainment and that background factors become relatively less important as children progress through junior school. Reading progress during primary school was related to socio-economic factors (social class background, low income and stage of fluency in English). Different progress rates in reading have been identified for different ethnic groups, with pupils from Caribbean background progressing the least.

According to Sammons, (*op. cit.*) when disadvantaged pupils make less progress in relation to their comparison groups, it is implied that the gap in reading attainment between working class children, those of low income, those lacking fluency in English and those of Caribbean background increased over time. In contrast, positive progress

rates should be manifested for the outcomes of disadvantaged pupil groups to converge with the majority group's outcomes. According to Thomas and Collier (1997, p. 67) minority pupils with limited proficiency in English '*must make more progress with each year of school than the typical native speaker makes to ever close the academic achievement gap on school tests*'. Yet, '*It remains important to examine the data on the attainment of different ethnic groups if we are to understand the complex situation and identify possible structural and institutional inequalities within the education system itself*' (Strand, 1999, p. 198).

However some School Effectiveness studies have not understood Thomas and Collier's statement; Haque and Bell (2001, p. 367), who conducted a school effectiveness study in secondary education, in the UK, wrote:

'Some of the lower performing Bangladeshi pupils at Key Stage 3 made considerable progress, since after two years they probably overcame the negative impacts of the re-cency of immigration. What is less expected is to find higher performing Bangladeshi pupils making much less improvement. It is difficult to pinpoint why this might be the case...'

Thomas and Collier's (*op. cit.*) statement and the fact that Sammons (1995) identified the fact that fully fluent minority pupils made less progress than pupils with limited competence in English indicate that there is a relationship between attainment and progress for minority ethnic groups. Ceiling effects restrain initially highly attaining pupil groups from demonstrating positive progress rates. It may prove to be difficult for pupil

groups characterized as above average to be positioned even higher in relation to their counterparts in a study conducted according to school effectiveness design.

Furthermore it is possible that foreign pupils may acquire the language of the host country at different rates as beginners and as intermediate and/or advanced learners of the second language.

Thomas's (1995) study investigated factors outside the control of the school that had an impact on pupil attainment (at the primary level as well as at the secondary level) in order to take into account of these factors in the analysis of school examination results. Thomas reported that pupils' socio-economic status, gender, age, ethnicity, home language and whether they had special educational needs had a significant impact on their later attainment. Apart from these individual pupils' characteristic, a contextual factor, namely the percentage of pupils in each school entitled to educational benefits was found to have an impact on the average attainment of pupils in each school. The study identified a substantial difference in attainment adjusted for pupils' background factors for different schools.

The finding of significantly lower final attainment score or underachievement for ethnic minority pupils in relation to the majority pupils, detected in previous studies (for example, in Swann, 1985) was replicated in the previous School Effectiveness studies. Swann's (*op. cit.*) study did not employ a school effectiveness design. Since the above School Effectiveness studies identified attainment differences between pupil groups, they have common constituent elements with another large-scale study entitled 'The

School Effect' (Smith and Tomlinson, 1989), which highlighted the school's role in promoting pupils' educational outcomes.

According to Teddlie et al. (2000, p. 84) recent multilevel re-analyses of data from the Junior School Project indicated that 14-15 per cent of the total variance in reading and in mathematics achievement at year 5 of the UK junior school, after adjusting for pupils' initial attainment score was due to school effects. It was confirmed that the school attended makes a difference to the final attainment score of a given pupil.

Tizard et al. (1988) investigated the attainment and progress of pupils in the infant school. In total 205 native and minority pupils in 33 multiethnic schools were tested at ages 4 and 7, at the beginning and at the end of the infant school. The study reported a significant interaction between ethnic group and sex. Caribbean girls made the most and Caribbean boys the least progress in reading and writing. White boys made the most progress in mathematics. Plewis (1991) followed up this sample at the end of Key Stage 2 (primary school) and identified similar differences between groups in reading defined by ethnic group membership and sex. Caribbean girls attained the highest and Caribbean boys the lowest. In mathematics there was some evidence that white girls caught up with white boys, but that Caribbean boys fell even further behind.

As part of the national evaluation of the Key Stage 1 Schagen (1994) collected assessments in English, mathematics and science. By using multilevel modelling Schagen explored the relationship between pupils' attainment in these subjects, their background characteristics and contextual characteristics of schools. He found that in terms of absolute attainment girls performed better than boys in English, and that there was a ten-

dency for ethnic minority groups to perform less well, with the exception of a significant positive effect in science for non white children who could not be categorised as black or Asian (which were coded as 'other'). English not being the first language had a significant impact in the English and science attainment targets. The number of terms in infant school had a positive effect on the mathematics and English attainment targets, and having attended nursery school had a positive effect on pupils' attainment in three out of four attainment targets. Schagen found no significant effects of most school or class level variables, including size and structure of class.

In the Netherlands Brandsma (1993, cited in Scheerens and Bosker, 1997), also investigated differences in compensatory potential of schools defined with respect to differences in pupil background characteristics such as socio-economic factors. Schools with a steeper regression of achievement on such background characteristics were seen as having a lower degree of compensatory potential than schools with a flatter regression slope. Among the variables examined, frequent evaluation and frequency of cooperation among teaching staff were found to promote pupils' outcomes.

Hill and Rowe (1996) carried out a longitudinal study of elementary schools during three years in Australia. This study splitted the total variation into three levels, (students within classrooms within schools) and concluded that classroom effects were very large and that the school effects were relatively small when the model adjusted for classroom participation.

Sammons (1995) examined over nine years the size and stability of gender, ethnic and socio-economic differences in students' educational achievement and progress rates.

Pupils were tested in reading and mathematics at primary school (year 3 and 5), at secondary transfer (year 6) and at the end of secondary education (year eleven). The net effects of these pupil background characteristics were estimated whilst controlling for the impact of other background factors and school membership.

Differences in attainment were identified for pupils from lower social class background throughout the testing occasions. Gender differences on attainment were not significant to start with, but they became significant on all the remaining testing occasions. Boys and pupils from lower social class progressed less than girls and pupils from higher social class respectively, a finding that was replicated throughout the junior school and secondary education. When certain groups make less progress in relation to their comparison groups, attainment differences increase over time.

Pupils who at entry to primary school were identified as having incomplete fluency in English had lower attainment in relation to pupils fluent in English in both subjects on all testing occasions, even up to university entry examination. This can be related to the fact that pupils with incomplete fluency in English were found to make less progress during the junior years in relation to fluent pupils. During the junior years differences in terms of progress were identified only for the Caribbean group who made less progress in English in relation to the English, Scottish Welsh and Irish group. In mathematics however, positive progress rates were identified for pupils from other ethnic backgrounds in relation to the ESWI group. The study demonstrated the complex nature of interactions between background factors and attainment over time and pointed to the value of examining both relative progress as well as absolute attainment differences at

any given age. The overall impact of background factors varied for different cognitive areas, being stronger for reading than for mathematics.

Sammons et al. (1997) highlighted the importance of background factors such as gender, level of income, level of fluency in English, on pupils' English attainment at the end of Key Stage 1, in comparison to pupils' attainment in mathematics and in science. Lack of fluency in English, socio-economic disadvantage (measured by eligibility in free school meals), term of birth and gender (boys doing less well) were significant predictors of pupils' final performance. Sammons showed that the magnitude of the individual effects was stronger for English than for mathematics and science and highlighted the importance of undertaking separate analyses for different subjects instead of using an overall composite measure of attainment. Sammons et al. (*op. cit.*) did not employ progress data, but accounted for intake differences through adjusting for individual pupils' characteristics.

Strand (1998) used multilevel modelling techniques to explore the results from performance tables in one local education authority with a fairer data analysis. The results revealed significant differences between schools' raw results as given in the performance tables and fair comparisons of their effectiveness in relation to pupils' ability based on an abstract reasoning test. The analysis indicated significant variation between schools in their effectiveness in different subject areas (English, mathematics and science) and provided evidence of underachievement among some groups of pupils. Bilingual pupils needing English language support, and pupils entitled to free school meals achieved lower KS2 results than would be expected from their reasoning ability in all three subjects and overall, based on an average test score. Pupils of Caribbean origin

achieved lower KS2 results than would be expected from their reasoning ability in mathematics, in science and according to the average test score. Boys achieved lower KS2 results than would be expected from their reasoning ability in language and in mathematics only.

Strand (1997, 1999) conducted studies related to the issue of equal educational opportunities of pupils between the ages of 4 and 7, and the value-added of the school. He explored the associations between ethnic group, sex, economic disadvantage and school attended on pupils' progress during their earliest years in school. Strand (1997) examined pupils' performance on national curriculum and Key Stage 1 assessments, and showed that different patterns of attainment and progress differences emerged for different groups of pupils. Girls had higher baseline attainment and also made more progress than boys during KS1, increasing the size of the gender gap in their attainment. Pupils entitled to free school meals (which is a proxy for social class) started with lower attainment and made less progress thus falling further behind their peers during the course of KS1. In contrast pupils with English as a second language (ESL) caught up with their monolingual English-speaking pupils (they made positive progress rates in relation to majority pupils). Pupils with more than three terms in early education and those with more than one term but less than three terms had higher attainment than those with one term or less. School compositional effects were also noted: pupils made on average more progress in schools with a high proportion of girls, and with high average baseline attainment score and less progress in schools with a high proportion of pupils entitled to free school meals, with a high proportion of ESL pupils.

Strand (1999) identified marked differences in pupil attainment associated with ethnic group, sex and economic disadvantage, both at baseline and at the end of Key Stage 1. In general, differences between groups of pupils tended to increase rather than to decrease over time. Pupils entitled to free school meals (FSM) made less progress in all subjects than pupils not entitled to FSM, girls made more progress than boys in reading and writing but less progress in mathematics and Caribbean pupils made less progress and Chinese pupils made more progress than English, Scottish, Welsh and Irish pupils.

Strand (*op. cit.*) tested the existence of differential effects in relation to ethnic group, sex and economic disadvantage and found no evidence of these effects. Strand (*op. cit.*) highlighted the need to investigate systematically interactions between ethnic group, gender and entitlement to free school meals (which in Strand's (*op. cit.*) study served as proxy variable for social class). Strand found that boys from certain minority ethnic groups and ESWNI pupils from lower social class background made significantly lower progress than expected.

These results are in line with the results of Gillborn and Gipps (1996) who although not adhering to the school effectiveness paradigm, found out that the effects of social class are interlinked with the effects of ethnicity and the effects of gender. Other authors also suggest that pupil background factors should be examined together as they have a joint effect on pupils' educational outcomes. According to Gilborn and Mizra (2000) ethnicity and social class effects are interlinked. Ethnicity and social class have a joint impact on pupils' outcomes. Meighen (1997) affirmed that the impact of variables representing pupils' individual social characteristics should not be assessed in isolation, but jointly, *'as social class, ethnicity and region, are intertwined with gender'*.

Strand (1997) Schagen (1994) and Thomas (1995) reported lower baseline attainment and positive progress rates for pupils with English as a Second Language in relation to native pupils for whom English was their first language. These authors also reported that eventually the former pupils caught up at the end of Key Stage 1. Strand (*op. cit.*) found that the condition of catching up is pupils spending three years in the host country's school by the age of seven. On the other hand, Sammons (1995) reported that pupils with incomplete fluency in English made less progress during the junior years (between the 3rd and 5th grades) and therefore attainment differences remained. Given that the former three studies refer to early primary education it is possible that it is easier for young minority pupils to catch up at the early grades of primary school than later on.

Strand (1997, 1999) revealed that the particular primary school a pupil attends has a significant effect on their progress during the early years. However, Strand (1999) found that there was no evidence of significant differential school effectiveness in relation to ethnic group, gender, or economic disadvantage. This implied that the same schools that were more effective for pupils not suffering from any form of disadvantage were also most effective for Caribbean pupils, boys, or economically disadvantaged pupils. Nor was any evidence for differential school effectiveness in relation to ethnic group, sex or social class in primary schools identified in the studies of Brandsma & Knuver (1989), and Sammons et al. (1993).

Hutchison (1993) analysed data from one LEA in England. Pupils were tested at the ages of six, eight and 12 in reading and he conducted a study of primary school effectiveness. He tested the impact of gender, age, nursery school education and social class (operationalised by free school meals) along with school level and contextual variables

on pupils' attainment and progress rates and found that receiving free school meals was negatively correlated with attainment and with pupils' progress. Also pupils having English as a second language had lower attainment and lower progress rates than the majority group. Boys were found to have lower attainment than girls but seemed to be catching up with them, demonstrating small positive progress rates. No relation was identified for age, terms in infant school and nursery school attendance.

Other factors known to relate to poor educational attainment are poverty, overcrowding, single parent families, poor attendance and large numbers of siblings (Fogelman, 1983, cited in Hutchison, 1993, p. 31).

Given that poor educational achievement is correlated with unemployment and with future low income as the British IPPR Report (Glennester, 1998) has shown, it is crucial that proactive measures are taken to alleviate attainment and progress differences between specific pupil groups. Support teaching programmes categorically targeting minority pupils with limited competence in the second language in particular and/or disadvantaged pupils in general aim to alleviate attainment differences between specific pupil groups and their comparison groups. Over time intervention or support teaching programmes aim to reduce these attainment differences by creating positive progress rates between minority and majority or between pupils from low and pupils from higher social class background.

The current study falls within the 'equal educational opportunities' research tradition, as one main goal is to assess the effect of what are sometimes seen as '*social determinants of achievement: gender, social class and race*' (Mortimore, Sammons, & Tho-

mas, 1994, o. 316) on pupils' final attainment scores at the end of primary school and on pupils' progress rates during this final year.

1. 2. 2. Describing the remaining four orientations within school effectiveness research

Research into education production function aims to establish a relationship or function between a variety of inputs such as family background influences, influences of peers, school inputs, innate abilities of students and a variety of student outcomes. *“The emphasis is on assessing school input-output relations that appear to be maximizing educational outcomes. The standard constrained-maximum model is formulated for the schools where output is reflected by a verbal achievement measure; inputs are composed of student characteristics, personnel attributes, facilities, and organizational variables”* (Levin, 1971, in McLellan et al., 2003). This function would accurately explain how a change in school inputs would affect school outputs. Measurable characteristics are used as inputs, such as the teacher-pupil relationship, teacher training, teacher experience, teachers' salaries and expenditure per pupil (Scheerens and Bosker, 1997, p. 143). The assumption behind education production function models is that increased inputs would bring increased students' outcomes.

Often studies dealing with the evaluation of compensatory programmes assess the effectiveness of such programmes that is to say to what extent pupils' enrolment in these programmes leads to an increase of disadvantaged pupils' educational and social outcomes (Slavin et al., 1990). Some evaluative studies look even deeper; they try to disentangle the important characteristics of these programmes, associated with increased

progress rates for disadvantaged groups of pupils. Compensatory programmes manipulate school conditions in order to raise achievement levels of disadvantaged groups of pupils. Often such programmes group together pupils from more than one class or grade and are delivered at the school level. Pollack et al. (1988) examined the differences found in the elementary schools that participated in the effective school process scheme, from an equity perspective. This study revealed important organisational, instructional and cultural characteristics between the equity and non-equity schools. Equity schools were considered those that managed to have equitable achievement scores across all economic sub-groups of the school's population. Viewed from an instructional perspective, in such schools, there is a greater emphasis on changes in classroom strategies for the benefit of low achieving students, particularly in the use of manipulative, hands on interactive strategies (e.g. co-operative learning). Non-equity schools, on the other hand, tended to rely more on resource people to assist low achieving students and rarely made school-wide changes in instructional practices.

Studies on the effectiveness of teachers, classes and instructional processes fall into the process-product research design. These studies relate antecedent conditions at the school or at the classroom level (e.g. content covered, quality of instruction, as well as psychological variables such as learning aptitudes and motivation) to the progress that pupils made during a pre-defined time period. Pupils' progress is reflected on the progress shown by schools or classrooms during the period between the two measurements, which is related to intervening variables. A study adhering into the educational productivity paradigm is Walberg's (1984).

In a process product design the impact of intervening variables on pupils' progress can be estimated. *'School effectiveness researchers aim to ascertain whether differential resources, processes and organisational arrangements affect student outcomes and, if so, how'* (Stoll and Mortimore, 1997). School effectiveness aims *'to identify common characteristics of effective schools and to establish criteria of measuring school effectiveness'* (J. Fraser in B. Creemers, 1989). In the USA, Levine (1990) summarised the results from numerous School Effectiveness studies and established effective school characteristics regardless of pupils' socio-economic background. Levine (*op. cit.*) focused his attention on studies describing schools characterised as effective that do not differ greatly in achievement from other schools of similar socio-economic composition, not demonstrating these outstanding results.

However, *'the causal status of the relationships found between school characteristics and effect measures is relatively small because of the correlative nature of the research, a lack of theory and insufficiently sharp-edged conceptualisation'* (Scheerens, 1992, p. 76). In Greece a PhD thesis adhering to the methodology of SER² that employed a process – product design has been carried out by Verdis (2002) in 375 upper secondary schools in Athens. Although the study identified significant school effects, the intra-school correlation was relatively small; it ranged between 2% and 10%. The study tested the impact of a number of statistical constructs (factors), namely 'teachers' responsiveness', 'students' academic self-image', 'principal's effectiveness', and 'collegiality among teachers', which were derived from exploratory factor analysis on pupils' progress. The study found that pupils' socio-economic status, age and sex had an im-

2 SER stands for School Effectiveness Research.

pact on pupils' final attainment and progress rates. The study also tested the impact of a number of school – level variables and found that large secondary schools (lykeia) had better results than small schools and that private lykeia had better results than state lykeia.

1. 3. School and classroom effects

Attending a given school or a classroom might have an effect on pupils' outcomes. The school effect can be described as *the relative impact of the school* in promoting pupils' educational outcomes. Classroom effects can be described as the relative impact of the classroom in promoting pupils' outcomes. School effects refer to differences between schools, while classroom effects refer to differences between classrooms within schools. School effects imply that there is significant variation at the school level when the total variability is partitioned between the school, the classroom and the pupil level. This hypothesis can be tested according to a range of absolute attainment and progress models. Among other studies Gray et al. (1995) and Thomas et al. (1997) investigated the existence and estimated the magnitude of school effects. If school effects are identified, then the magnitude of school effects is estimated through the calculation of the Variance Partition Coefficient or VPC (Goldstein, 2002, p. 3), which shows the relative importance of the school level variance compared to the total variance. If classroom effects only are identified, then the VPC shows the relative importance of the classroom level variance compared to the total variance. In the case that a model with three levels is fitted then the total variance is the sum of variances at the school level, the classroom level and the pupil level. In the case that a model with two levels is fitted then the total

variance is the sum of variances allocated at the higher level, which stands either for the school level or for the classroom level, and at the pupil level.

Teddlie and Stringfield (1993, p. 25), who conducted the Louisiana School Effectiveness study in the USA, concluded that *'75% of the variation in individual student achievement could be linked to student characteristics, 12% could be linked to teachers, and fully 13% could be linked to differences between schools'*. There is a growing recognition that school effects research must take into account what happens in classrooms. As several researchers have noted (e.g. Creemers, 1992; Reynolds, 1992; Scheerens, Vermulen & Pelgrum, 1989), a large percentage of the variation among schools is due to classroom variation.

Scheerens et al. (1989, cited in Scheerens and Bosker, 1997) in their international review of School Effectiveness studies, looked at the size of unadjusted school and classroom effects as well as effects adjusted for pupils' social class. They concluded that internationally the contribution of school-level variance to the total variance of students' outcome measures is approximately 10 per cent.

Table 1. 1: Variance Partition Coefficients for classrooms and schools for different countries from Scheerens et al. (1989)

<i>Country</i>	<i>Classroom variance component</i>	<i>School Variance Component</i>
Canada	0.18	0.009
Finland	0.45	0.002
France	0.17	0.006
New Zealand	0.45	0.01
Scotland	0.34	0.12
Sweden	0.45	0.00
USA	0.46	0.10

Several countries (USA, Sweden, New Zealand and Finland) showed relatively small differences between schools but large differences between classes within schools.

In Canada and in France relatively small differences both between schools and between classrooms within schools are shown.

Luyten (2003, p. 35) employs a measure of resemblance between classrooms within schools (ρ), which can be estimated by dividing the school-level variance by the sum of the classroom-level and school-level variance. This statistic is used for studies that partition the total variability into three levels.

In the formula:

$$\rho = \frac{\text{school-level variance}}{\text{school-level variance} + \text{classroom-level variance}}$$

ρ is the ratio of school level by the sum of school and classroom level variances (ρ refers only to variance terms above the level 1). Hence $1-\rho$ is the ratio of classroom-level by the sum of school-level and classroom-level variances and it is indicative of the importance of classroom effect.

Luyten (2003) reviewed studies in which pupils' outcomes had been adjusted for both prior attainment and other relevant intake variables such as family socio-economic background, ethnicity, gender and class composition.

Luyten's (2003) compares pupils' outcomes derived from parallel classes, that is outcomes derived from classrooms serving pupils who are enrolled in the same grade in the same school, but in different classrooms. Comparisons between parallel classes are

more indicative of classroom effects, as classroom effects are not confounded by the fact that one is comparing different age groups and / or different curriculum contents.

Research papers that provide information on the relative size of school and teacher variance turn out to be very rare.

Table 1. 2: Results of ρ and school-level variances for parallel classes in primary schools

Study	Results for language:		Results for mathematics:	
	ρ	School-level variance	ρ	School-level variance
Mortimore et al., 1988	0.40 (median ρ)	6.1%	0.21 (median ρ)	5.1%
Bosker, 1991 The Netherlands	0.96	10.5%	0.69	14.3%
Hill & Row, 1996 Australia, Victoria	0.16	8.2%	0.9	5.4%
Median ρ language 0.40	Median school-level variance language 8.2%			
Median ρ mathematics 0.21	Median school-level variance mathematics 5.4%			
Median ρ across subjects 0.31	Median school-level variance both subjects 7.2%			

For language, the median ρ across studies is 0.40, which implies that teacher effects are about one and half times as large as school effects. For mathematics, the median ρ across studies is smaller, namely 0.21. In other words, for primary schools mathematics teacher effects are nearly four times as large as school effects. In the Netherlands school effects are much greater than classroom effects (Bosker, 1991) while in Australia classroom effects predominate (Hill and Row, 1996).

The VPC(s) derived from the analysis carried out in the current study refer to the partition of classroom level variance in relation to total variance (levels 2 and 1), whereas in Luyten's analysis $\rho(s)$ compare school-level and classroom-level variances.

Examples of studies that identified school effects in primary schools by using a multi-level design were in the UK Mortimore et al. (1988), Tizard et al. (1988), Thomas (1995), Sammons (1995), Plewis (1991), Strand (1997, 1999), Bondi (1991), Tymms et al. (1997), in the Netherlands Brandsma (1993, cited in Scheerens and Bosker, 1997, and in the USA Teddlie et al. (2000). Tizard et al. (*op. cit.*), although they did not study school effects in a systematic way, reported significant differences between schools in the amount of progress made by pupils between four and seven years of age.

In Greece, school effects were identified in secondary schools only (Verdis, 2002).

1.4. Differential effectiveness for different student groups

Differential school effects concern the existence of systematic differences in attainment between schools (or classrooms) for different pupil groups (e.g. pupils with different levels of prior attainment (low achievers versus high achievers) or pupils with different individual characteristics, (such as ethnicity, social class and gender), once the average differences between groups have been accounted for (Sammons, 1996). The analysis of differential and contextual school effects is particularly helpful in order to highlight equity issues among pupils at the school or at the classroom level, which are particularly important for policies aiming to address the equal educational opportunities issue.

Willms (1992, cited in Teddlie et al., 2000, p. 127) argues that among individual pupils' characteristics, measures of socio-economic status are particularly important with respect to differential school effects.

If differential effects exist, this might mean that some schools/classrooms implicitly or explicitly favour certain groups of pupils (e.g. high achievers, ethnic majority pupils, etc.). In cases where differential effects exist, the impact of variables on which differential effects are found is not stable, but varies from one school/classroom to the next. The differential effectiveness hypothesis serves to disentangle issues of equity. *'One of the major issues in educational research concerns the effectiveness of schools in relationship to pupils' background characteristics, especially the question whether schools can do better than other schools for the disadvantaged'* (Striengfield et al., 1997 in Creemers et al., 2002).

Thomas et al. (1994, in MacBeath, 1999, p. 2) found that *'schools which obtained higher than average results for the most able (so called 'effective schools') were ineffective for the least able'*. This finding underlines the point that on top of assessing schools or classrooms based on their overall effectiveness (in terms of progress, or adjusted progress) one should go deeper and ask the question whether schools or classrooms are equally effective for different categories of pupils enrolled in them.

Evidence of differential school effects was found in the studies of Smith and Tomlinson's (1989), Nuttall et al. (1989) and in the Hampshire Project (Goldstein et al., 2000). According to Smith and Tomlinson's study students with different levels of prior attainment performed differently in different schools. In particular, differences in Eng-

lish between schools were greater for students with above-average than for students with below-average second year reading scores. Nuttall et al.'s study found that while some schools were narrowing the gap between students of high and low attainment on entry, other schools did not. Nuttall et al.'s study (1989) highlighted the existence of differential school effects for students of different ethnic backgrounds. This study reported '*within-school*' differences in effectiveness between Caribbean and English/Scottish / Welsh (ESW) students' and commented that '*other ethnic differences vary across schools even more than the Caribbean-ESW differences*'. Goldstein et al. (2000, p. 5) also found evidence of differential effectiveness '*with schools differing in terms of the progress made by pupils with different intake (KS1) achievements*'.

As far as differential effectiveness and socio-economic status are concerned, Mortimore et al. (1988a) and the reanalysis by Sammons et al. (1993) found evidence of significant differential school effectiveness among English primary schools for specific ethnic groups. Thomas et al. (1995) found that in their study, although all categories of pupils progressed in effective schools, some groups (those not disadvantaged) performed especially well. In summary, whilst disadvantaged groups did better in more effective than in less effective schools, the gap in achievement between the disadvantaged and the advantaged pupils increased within the more effective schools. Another British study, Strand (1999) tested the differential effectiveness hypothesis and concluded that from baseline to the end of Key Stage One, there was no evidence of differential school effectiveness in relation to ethnic group, sex or economic disadvantage. Brandsma and Knuver (1989) tested the existence of differential effects in primary schools in the Netherlands for arithmetic progress and for language and found limited differential ef-

fects of schools in relation to pupils' prior attainment for language only. They tested the existence of differential effects in relation to pupils' gender for mathematics and language and found only evidence of differential school effects for the Dutch language. In contrast, they found no evidence of the existence of differential school effects according to ethnic groups. Sammons et al. (1993) investigated whether differential school effects existed in primary schools but found no evidence for differential school effectiveness in relation to ethnic group, sex or social class.

The issue of differential effectiveness is important for the current study because differential progress rates associated with ethnicity can reflect the adoption or not of different support teaching programmes and different priorities placed onto the equity dimension by different schools / classrooms. Schools or classrooms that create more progress for initially low achieving as well as for disadvantaged groups of pupils put a stronger emphasis on the equity dimension of schooling, in relation to schools or classrooms that boost more the outcomes of initially high achieving pupils. The adoption of compensatory education schemes at the school level, appropriate pedagogy in the mainstream and a supportive school and classroom climate are among the factors that can contribute to equitable outcomes for disadvantaged pupil groups.

If support education is implemented in a coherent way throughout the schools that take place in the study, schools having support teaching policies may have less steep slopes associated with prior attainment scores than schools without explicit support teaching policies. If this scenario is true, then schools providing support teaching would have produced more progress for low achieving groups of pupils than schools without a support teaching policy.

1. 5. Identifying links between School Effectiveness and School Improvement

Recently there have been attempts to link the educational fields related to School Effectiveness and School Improvement. SER can be used as an educational monitoring device at districts' disposal. SER can identify low achieving classrooms and schools, or classrooms/schools where discrepancies in attainment and/or in progress between minority and majority ethnic groups or between pupils from higher and lower social class occur. Finally SER can identify lessons or curriculum areas where significant weaknesses are identified for these classrooms/schools, so that steps can be taken to alleviate them. These steps may involve drawing the profile of classrooms or schools on pupils' attainment and progress rates on every curriculum section tested. Gray and Wilcox (1995) state that SER findings can inform policy development, and be subsequently translated into educational practice. If SER can identify factors related to increased pupil attainment, then classrooms/schools can modify their policies and practices to include these factors.

Stoll (1996, p. 55) perceives school improvement to be the *'ultimate test of the theories produced by school effectiveness researchers'*.

Findings derived from both SER and Educational Research in general can be tried out by classrooms/schools for school improvement purposes. However, in order to ensure that an innovation would work in the specific classroom/school setting(s), districts and/or schools may undertake evaluations of the educational innovation adopted. These evaluations involve pupils' testing in core subjects twice, before and after this innovation takes place; then pupils' progress during the time the innovation was implemented

can be compared with the progress of pupils of the same grade in settings that did not implement the innovation. One of the core questions of such an evaluation would be '*What might work for whom and in what circumstances*' (Pawson & Tilley, 1997). So a range of data may be collected and multi-method analysis may be undertaken (including multi-level modelling) to answer these questions.

Hopkins et al. (1994, p. 66) state that '*students' outcomes are the fundamental goal for educational reform*'. Pupils' progress rates and final attainment scores are among the best indicators of the success of a particular innovation.

Van Velzen et al. (1985, p. 48) described the School Improvement process as: '*A systematic sustained effort aimed at change in learning conditions and other related internal conditions in one or more schools with the ultimate aim of accomplishing educational goals more effectively.*'

According to Stoll (1999, p. 504) '*if all students are to realise their potential, then their schools must also realise their potential*'.

Creemers and Reezigt (1997) stated that a distinction is often made between theory and research on the one hand, outlined in School Effectiveness studies, and practice and policy making on the other, outlined in school improvement studies.

Despite differences in approach and orientation of the two fields, their findings are, for the most part, complementary. While the school effectiveness factors represent a static picture of what one would expect to find in an effective school, school improvement 'factors' or conditions can be perceived as part of an overall change strategy or route.

Also, school improvement might give more insight into the strategies to be employed by the school successfully in the direction of effectiveness (Maughan et al., 1990).

Reynolds et al. (2000a, p. 206) suggested that there are clear interests involved in maximising the relationships between school effectiveness and school improvement, since effectiveness research can provide the knowledge base concerning what to do, and the improvement community the vehicle within which the knowledge base sits.

The reorganisation of support for low achieving pupils appears among the guidelines for school improvement created by the English Office of Standards in Education (OFSTED (1994c) in Gray and Wilcox (1995, p. 246)). Other guidelines appearing in the same report are the *greater use of assessment* and the continuous development of transitional arrangements between primary and secondary school. These organisational arrangements can be perceived as preparation of reports accompanying minority and low achieving pupils in their transfer to secondary school, or when they are enrolled in another primary school. These reports may refer to support teaching that minority pupils and low achievers have been exposed to so far and on the curriculum covered in the support class. According to OFSTED (1994c) another area where school effectiveness results can inform school improvement is that schools can monitor their work on the basis of SER findings, so as to keep momentum and standards up. The need for the school to monitor its work is also referred to in the initial 5-factors school effectiveness model by Edmonds (1979). Although all schools/ classrooms can improve their practices and policies, school improvement is particularly crucial for classrooms and schools performing at a below-average level.

Mortimore (1998, p. 330) argued that school effectiveness and school improvement research highlighted the relationship between disadvantage and school improvement.

Mortimore (*op. cit.*, p. 332) believed in the necessity for schools to nurture the disadvantaged pupils in order to bring them up to the same starting line, from which more advantaged pupils begin their schooling. Mortimore also argued for the need to establish early prevention or early intervention programmes in every school. He also firmly believed that the principals needed to find new ways of supporting the disadvantaged (p. 334).

Harris (2001, p. 16) judged the development of school improvement strategies, which have an impact simultaneously and consistently at whole school, teacher and pupil level, as necessary. The creation of support schemes for low achieving pupils has the characteristics described in Harris (2001): Support schemes are organised in the school, but they have a positive impact on the quality of teaching and learning in the mainstream classroom (it becomes easier for the mainstream class teacher to teach a more homogeneous classroom); and they have a direct positive impact on the learning of individual pupils, as they enable both native and minority pupils with low attainment to understand the curriculum in the mainstream class.

Macbeath and Mortimore (2001, p. 38) described three themes as necessary preconditions for school improvement: School Ethos, the process of Development Planning and a particular focus on Learning and Teaching.

School improvement '*has emphasised the importance of focusing attention at the student level and of improving teaching and learning conditions within the classroom*' (Hopkins et al., 1997).

1. 6. Transferability of constructs and variables

A lesson that can be drawn from School Effectiveness studies carried out internationally is that school effectiveness researchers should proceed with extra caution in adopting constructs³ and variables from the literature. Existing school effectiveness instruments may measure constructs (e.g. classroom climate) operationalised in other educational systems. Therefore these existing instruments may measure constructs and variables of foreign educational systems instead of the Greek system. One big difference is that the Greek educational system is centralised (Kavouri, 1996), whereas the English and American ones are decentralised. More specifically, individual questions and items included in these instruments may not reflect the realities of the Greek schools and educational system, whereas other items, which are not included in these existing instruments, may better reflect the Greek reality. Therefore once amended, it would be wise to pilot the new instruments in the Greek schools, before they are used on a large-scale basis in a school effectiveness study. These instruments have to be sensitive and reflect the Greek organisational arrangements. According to Scheerens and Bosker (*op. cit.*), if process variables do not capture the whole range of variability associated with organisational processes at the school level or with classroom processes, small effects are obtained on the impact of these process variables.



1. 7. Case studies based on a SER design

By examining the pedagogical approaches and teaching practices occurring in classrooms performing at an above-average level, in comparison to classrooms performing at an average level, and to classrooms performing at below-average level, one may begin to shape an idea about '*which instructional practices are working*' within the given educational settings. Such a design was adopted by the 'study of Louisiana schools' in the USA (Teddle, and Stringfield, 1993). Instructional and organisational practices adopted in schools with above-average, average and below-average performance were examined qualitatively, in a sample of schools falling within each one of the three categories. The educational outcomes of each one of the schools participating in the study were assessed in a quantitative way. This study adopted a mixed design and more specifically one that adheres to the process – product research design. Classroom/school processes were associated with differences in progress between classrooms and/or between schools. Teddle and Stringfield (1993) reported that classroom practices in ineffective schools were characterized by lower rates of time on task, less teacher presentation of new material, lower rates of teacher communication of high academic expectations, fewer instances of positive reinforcement, more classroom interruptions, more discipline problems, and a classroom ambience generally rated as less friendly.

³ These constructs are theoretical, non-observable, inferred variables (such as school or classroom climate), which are operationally defined through the measurement of other variables and their theoretically expected relations.

Before the Louisiana study, Weber (1971) in Good and Brophy (1986) examined the characteristics of high achieving urban, inner city schools in connection with teaching reading. More specifically he identified school factors or characteristics distinguishing high achieving schools from other schools. He cited eight successful school characteristics: strong leadership, high expectations, good atmosphere, placement of extra emphasis on reading, employment of additional reading personnel, use of phonics, individualisation and careful evaluation.

1. 8. Other issues:

Other issues preoccupying school effectiveness researchers as described by Sammons (1996) are:

- The size and significance of effects
- Consistency across outcomes
- Stability over time
- The long-term effects of schools
- Differential effectiveness for different student groups

Reynolds (2000) refers to two other issues:

- The processes within schools
- The possible 'context specificity' of these process factors.

1.8.1. Size and significance of effects

Elliot (1996) claimed that school effects are unimportant because they account for only 8-13% of variation in student outcome measures. However, other SER studies have demonstrated that 40-50% of the total variation in students' outcome measures is accounted for by various factors considered jointly (Sammons & Reynolds, 1997). Therefore the school effects represent a sizeable component of what can be accounted for in terms of pupils' outcomes.

1. 8. 2. Consistency across outcomes

In contrast to older School Effectiveness studies like Rutter's (1983), that claimed that schools were equally effective across a range of outcomes, many contemporary School Effectiveness studies, (e.g. Thomas et al., 1996, Sammons, 1994, Thomas and Mortimore, 1996), found that schools are not consistently effective across different subject areas. *'In terms of effectiveness the widely held view is that only a small number of schools are excellent across the board'* (Reynolds, 1996, p. 174).

'Few schools are consistently effective across a range of subjects. Some studies have cautioned against an over-reliance on aggregate exam scores in assessing school effectiveness' (Sammons et al., 1997; Thomas et al., 1997 in Smyth, 1999).

Consistency between pupils' outcomes in different subjects is measured by examining the correlation coefficients of the school residuals obtained in these two subjects, as in the study by Thomas, Sammons and Mortimore (1995). School Effectiveness studies

looking into the consistency issue examine whether '*the rank order of schools according to their effects in the mathematics domain is congruent with their rank order in the language domain*' (Scheerens and Bosker, 1997).

Goldstein et al. (1993) found a correlation coefficient in the low range ($r = 0.3$) between examination results obtained by schools in English and in mathematics in the UK. Also in the UK, Sammons et al. (1993) reported a consistency estimate between mathematics and reading of 0.62.

The Mandeville (1988) study investigated the correlation between the effects on mathematics and reading. The results indicated fairly substantial consistency (the correlation coefficient between the two subjects approximating 0.70), a figure in line with the correlation coefficient reported by Bosker's (1990) study of elementary schools in the Netherlands ($r = 0.73$), between language and mathematics. Van Batenbourg (1990, cited in Scheerens and Bosker (1997)) reports correlations in the range of 0.80-0.90 for gross school effects. Scheerens and Bosker (1997) who summarised the findings from studies in the consistency of school effects across subjects (basically mathematics and reading) in primary education from research carried out in the USA, the UK, Australia and the Netherlands estimated that the median consistency estimate across all studies was 0.62. This was exactly the same estimate that Sammons et al. (1993) found for 10-year old students in the UK.

In addition, some studies examined whether schools are consistently effective across a range of cognitive and social / affective outcomes (Rutter et al., 1979, Mortimore et al., 1988 in Sammons, 1996). According to Teddlie et al. (2000, p. 116) '*some researchers*

in the US have usually confined these dependent measures to academic achievement with some exceptions (e.g. Kochan, Tashakkori, and Teddlie, 1996; Teddlie and Stringfield, 1993), while researchers in the UK have been more likely to consider other types of outcomes'. Outcomes other than academic achievement are attendance, attitudes, behaviour and self-concept (Sammons et al., 1996a).

1. 8. 3. Stability over time

The stability question is crucial for the validity of the conclusions derived from School Effectiveness studies. Rowe (2000, p. 80) suggests that *'schools or teachers within these schools should not be judged by a simple cohort of students, but rather on their performance over time'*.

The stability issue is crucial for the following reasons: No conclusions can be reached about the effectiveness of particular classrooms or schools or about effective classroom or school correlates in case there is instability in the way classrooms or schools are characterised as above average, average or below average, across several testing occasions. Also in case there is instability in the relative rankings of classrooms or schools, one is uncertain about the conclusions reached through a given School Effectiveness study. Unstable classroom or school estimates imply that correlation coefficients between classroom/school residuals derived from successive replications of the School Effectiveness study are relatively small. These classroom or school residuals would have been derived from studies based on the same sample of classrooms or schools. For example, instability in classroom/school residuals may occur as a result of some class-

rooms or schools performing at an above-average level in the first study, but then performing at an average or at a below-average level during the year that the (School Effectiveness) study was replicated.

The stability of school effects calculated between school/classroom estimates derived from different cohorts and grades varies in studies conducted in different countries. In the USA Mandeville (1988) addressed the issue of stability of school effects. He conducted a School Effectiveness study during 1-4 grades in primary schools and calculated the correlation coefficients between the school's rankings derived from different cohorts of pupils. These correlation coefficients ranged from 0.34 to 0.66.

In the Netherlands, Van Batenburg (1990, cited in Scheerens and Bosker, 1997) analysed various cohorts of students graduating from primary education, and found that the correlation between the rank orderings of schools per year cohort, varied between 0.78 and 0.83. Block and Hoeksma (1993) (in Scheerens and Bosker, 1997) conducted a similar study to Van Batenburg's, and concluded that the stability estimate for arithmetic (0.87) was somewhat higher than that for language (0.84) and reading (0.80). Luyten (1994) concluded that '*schools produce fairly stable results across years*' or that '*the general year effect turns out to be very modest*'.

In Britain, Nuttall et al. (1989) have identified instability of individual school estimates, whereas Goldstein (1987) has shown consistency in school outcomes over time.

As in some School Effectiveness studies there is instability in school estimates over time, school estimates from more than one replication of this School Effectiveness study are required in order to draw valid conclusions about *individual school effects*.

Scheerens and Bosker (1997, p. 95) firmly believed that schools are less stable in terms of effectiveness when the time interval is longer than one school year.

The stability of school estimates across different grades at the same time is another issue addressed in Grisay's (1996) study, involving 94 secondary schools in France. The school effects were assessed at the same time, in grades 3 and 5 of secondary school. It was found that for mathematics the gross school effects (derived from null models) had a correlation of 0.79, whereas for French language this correlation is 0.84. In terms of net school effects, (derived through the progress model), the results are less highly correlated: 0.42 for French language and 0.27 for mathematics.

These results show that even in cases where school effects exist, these results might differ across different grades belonging to the same school at the same time point.

According to Scheerens and Bosker (1997), grade specific variation may either point to teacher effects being more important than school effects, or to the test-specific variation that results as a function of using performance tests based on the curriculum of different grades.

1. 8. 4. Long term effects of schools

A study carried out by Sammons et al. (1995b) examined whether primary school effects continue to have an impact on students' attainment or progress rates, even when these students transfer to secondary school. These results suggested that the primary school attended exerts a long- term effect upon pupils' performance, even after controlling for their attainment at secondary transfer.

Goldstein and Sammons (1995) argue that the secondary school effectiveness model consistently over-estimates the effect of secondary school. They concluded that '*the usual quantitative procedures for estimating school effectiveness need to be augmented with careful measurements of all relevant prior performances including the effects of the primary schools attended*'.

This statement suggests that the future performance of foreign and repatriated pupils in secondary schools will inevitably be influenced by the primary school in which they were enrolled as well as by the type of support they have been subjected to during primary school.

'In a decade when educational policy prescriptions are travelling internationally with increasing frequency, it is unfortunate that school effectiveness researchers still seem locked into a mind-set that can handle only the patterns of intakes, processes and outcomes of their own cultures, rather than attempting to throw light on issues of possible context-specificity or universality' Reynolds et al. (2002, p. 5).

The authors believe that school effectiveness should move away from the 'one size fits all' tradition, as new findings of context specificity do not fit into the dominant five, seven or thirteen-factor model of SER. For example, research carried out in the Netherlands did not verify the importance of school leadership as a vital constituent part of effective schools. Reynolds et al. (2002, p. 5) implied that school effectiveness researchers sometimes did not highlight issues of context specificity for fear that if they did point to these issues, they might reduce the certainty in the findings of SER, reduce the status of SER and consequently their position within the research community.

'With effectiveness researchers now enjoying considerable growth in the attention given to them, in the research resources available to them and in the practitioner communities willingness to buy their services, a group of people who had historically been distrusted by their professional peers and marginalized within their own professional research communities might have been excused for not rushing to complicate their knowledge bases with any notions of context-specificity that might have made their official acceptance more problematic' (Reynolds et al., 2002, p. 13).

1. 9. Outcomes measured

In many school effectiveness studies the normative performance-based tests employed measure performance in basic skills only (language and mathematics); higher order skills were rarely included in tests. Levine (1992, p. 27) believed that higher-order skills should be sufficiently represented in the performance tests employed in SER studies; there should be more emphasis on reading comprehension and problem-solving in mathematics which should be represented in sections of the standardised tests employed, rather than only language mechanics and maths computation to be included in the tests; therefore newly developed assessment instruments should be introduced focusing on higher order learning and thinking skills (Guthrie, 1987; Cooper, 1989; Ivens and Koslin, 1989).

The impact of attending a certain school or classroom was not related to the promotion of social and affective outcomes (such as pupils' satisfaction, motivation, self-perception, socialisation etc.) nor was the impact of effective school correlates on social and affective outcomes frequently investigated. *'The need for multiple outcome measures has been stressed in every major review of the field since the early 1980s (e.g. Scheerens and Bosker, 1997; Good and Brophy, 1986; Rutter, 1983; Scheerens, 1992; Sammons et al., 1995; Teddlie and Stringfield, 1993 in Teddlie et al., 2002)'*. Only recently the International School Effectiveness Project (Teddlie et al., 2002) has measured the impact of schooling on affective outcomes. In many British studies in addition

to academic outcomes, '*other factors such as rates of attendance and rates of delinquency were incorporated*' (Creemers 1996, p. 39).

1. 10. Other School Effectiveness Criticisms

Pring (1996), (cited in Goldstein & Woodhouse, 2000) criticises some school effectiveness researchers for having accepted the politically acknowledged assumption that schools can and should be held responsible for economic and social improvement.

Goldstein & Woodhouse (2000) have summarised the following critiques against SER under the following broad headlines:

- Abuse by Government
- Oversimplification of the complex 'causalities' associated with schooling and side-tracking into focusing on 'league tables'
- That 'theory' in SER work is little more than reification of empirical relationships
- Too much SER is simply poor quality.

Goldstein and Woodhouse (2000, p. 10) contest the position of Sammons et al. (1996) that '*research should inform policy*'. Goldstein and Woodhouse (*op. cit.*) believe that there should be a clear separation of SER from governmental influence. They suggest that while SER may result in practical policies, its primary justification is to increase knowledge and understanding.

Gibson and Asthana (1998) believe that SER has reinforced government policies which are concerned with identifying schools as the sole agents of 'success' or 'failure' and

they illustrate another important issue, namely that not only will governments selectively use findings, they and their advisors also often simply fail to understand them.

Elliot (1996) criticises Sammons et al. (1995) for '*adopting a mechanistic methodology and an instrumental view of educational processes*'. He regards SER activities as '*ideological legitimisations of a socially coercive view of schooling*'.

Elliot (*op. cit.*) highlights the claims made by several SE⁴ researchers that they need to make their research relevant and practical, in the sense that it may be employed for a certain purpose such as for school improvement. He suggests that such a concern endears SER to government and can lead both to a superficiality of interpretation and a concentration on *pragmatics* rather than theory. However, Reynolds (1998, p. 20, cited in Thrupp, 2001, p. 447) believes that the fact that SER is more practically than theoretically oriented is an asset, rather than a drawback as SER, not wasting time in philosophical debates, can make rapid progress.

Thrupp (2001, p. 447) believes that this inherent pragmatism in SER limits the analysis and makes it prone to ideological capture. Yet, although this might have occurred in the UK, it does not necessarily imply that the same phenomenon will occur in other settings and cultures.

4 SE stands for school effectiveness

As far as the limitations of SER are concerned, there is no need for SER to endorse any political orientation, especially when the latter concerns accountability amongst schools. However, the use of SER for redistribution purposes in favour of the disadvantaged pupils is an act with a political character as well, aiming to deal with equity issues and to address the impact of social class and ethnicity on pupils' initial and final attainment scores and on their progress. Given that pupils' attainment at entry is correlated to pupils' social class membership, and that initial attainment score is the strongest predictor to pupils' final attainment score (Goldstein, 1997) in the absence of favourable support teaching arrangements or/and of additional measures taken in the mainstream class it is expected that disadvantaged pupils would attain a lower final attainment score on average than advantaged pupils. Goldstein (1998, p. 3) states that pupils' socio-economic group has a strong effect on pupils' progress during Key Stage 1 (first grades of primary school) in the UK. When pupils begin their schooling their initial attainment is already affected by pupils' social class membership. When SER analysts in order to estimate pupils' progress at a higher grade control for initial attainment score at entry, they parcel out the impact of pupils' social class at entry, which is correlated with their initial attainment score.

Adequate support schemes in schools would be able to address such educational inequalities. In the absence of positive discrimination arrangements, the impact of social class is likely to remain intact throughout pupils' schooling. A large-scale longitudinal study of primary schools carried out by Mortimore et al. (1988) found that *'no school reversed the usual 'within-school pattern of advantaged pupils performing better than the disadvantaged'*. What may be required from the school though, is that compensatory mechanisms are in place in order to alleviate the impact of social class,

tory mechanisms are in place in order to alleviate the impact of social class, or the impact of ethnicity, which in certain cases and for certain subjects become more pronounced the more years a disadvantaged pupil spends in the school. This is demonstrated in the analyses when the social class has a significant impact in terms of progress, which implies that disadvantaged pupils fall further and further behind from their advantaged peers. *'SER treats the class backgrounds of students as a given when of course they are not really given at all, they are socially constructed, and can be made worse or better through housing, health, employment, and taxation policies, all of which will therefore affect student achievement'* (Anyon, 1997, cited in Thrupp, 2001, p. 448). But as SER wants to highlight the point that the school attended can make a difference it should highlight the need for the impact of social class to be reduced throughout each pupil's schooling, basically through the school's efforts. Mortimore et al. (1997) suggest *'a continuing need for positive discrimination and the effective targeting of human and material resources'*. This can be achieved through a redistribution mechanism that the school would set, e.g. homework support, extra language tuition and assistance in mathematics, and to a lesser extent with the provision of extra-curricular or enrichment activities for all the pupils. Such an action may be justified on an *equity principle*. Briefly, the justification used by the educational system for setting up remedial classes and allocating additional funding to pupils with special educational needs applies also to pupils who face an additional disadvantage due to their inadequate competence in their second language, or due to their lower social class background. Pupils who come from a milieu providing them with fewer educational and cognitive experiences or stimuli should be endowed with extra resources in the school framework to

overcome educational disadvantage. Harvey and Klein (1989, p. 57 in Apple 1997, p. 254) describe '*the need to provide appropriate routes of access to the school's services that would allow everyone to avail themselves of existing educational treatments and benefits*'. Different strategies and facilities (e.g. differentiated support teaching programmes, individual educational plans, differentiated curriculum or materials to be additionally employed in the mainstream) can be justified for this purpose.

Teddlie (in Thrupp, 2001, p. 447) stressed that SER has long agreed that social class does have a major impact on student achievement and that SER has not ignored or downplayed the importance of class context, as often claimed by critics of SER. In support of this, he points to context factors being an important focus of SER, citing numerous studies over the last decade. The two empirical studies, *which mention the contextual impact of social class* are the studies of Hallinger and Murphy (1986) and Teddlie et al. (1989) (cited in Thrupp, 2001, p. 448). Thrupp (1997, 1999) argued that academic achievements of students in predominantly working class schools is likely to be depressed below their individual level of ability and social origin as a result of social processes within schools.

Mortimore (1998, p. 323) ascertains that SER has been concerned with questions of equity and as examples of studies that dealt with the issue of equity on pupils' educational outcomes and cites 'School Matters' (Mortimore et al., 1988a) and Sammons (1995). Mortimore (1998) firmly believes that the negative impact of social and economic disadvantage on students' educational opportunities is dealt with in SER on a constant basis (e.g. Mortimore, 1995d, 1996b; Mortimore & Goldstein, 1996; Sammons et al., 1994c, cited in Mortimore, 1998, p. 323). However, SER argues that fair comparisons

between schools/classrooms should be undertaken, controlling for pupil intake, or for other contextual characteristics of the school as well, so that only similar schools/classrooms in terms of pupils' socio-economic characteristics, pupils' initial attainment score and contextual characteristics, are compared. Such a practice may guarantee that fair comparisons between schools/classrooms are undertaken. This debate emphasises educational accountability of individual schools/classrooms so that good schools/classrooms take credit for their good results, and 'bad' schools/classrooms are pressed by the inspectorate to improve. The underlying idea behind this is that school/classroom comparisons should not be based on their pupil intake. On the other hand, SER has not been adequately involved with the districts' and schools' responsibility to redistribute educational outcomes for educationally and socially disadvantaged sections of the school population. In England such a provision is planned for schools belonging to education action zones (*DfEE, 1997:4* in Gamarnikow et al., 1999). Support teaching to low achieving groups including ethnic minorities may take place in individual schools, where these schools have an equal educational opportunities policy. Although the English Office for Standards in Education (OFSTED, 2000a) has adopted explicit criteria when judging schools referring to the performance of subgroups of pupils, a general approach on backing up minority pupils and pupils from lower social class who underachieve with support teaching is not currently in the educational agenda.

Sometimes variables employed in SER are rather proxies than elaborated/refined. For example, '*free school meals*' has been used as a proxy variable for social class in the

majority of School Effectiveness studies in UK and is characterised as crude by Thomas (1995). Instead a social class variable should be defined based on sociological theory.

Chitty (1997, p. 55) summarises in the following way the criticisms associated with SER:

- SER places too much emphasis on progressive school management as the dynamic of change;
- It fails to account fully for the characteristics of the educational system as a whole;
- It shows little regard for issues of social class;
- It has little or nothing to say about curriculum content and pedagogy.

SER emphasises the importance of school leadership, administration, school climate or collegiality amongst teachers but it puts less emphasis on specific classroom conditions. With few exceptions, among which are Creemers (1994, 1997) and researchers influenced by his work (e.g. Kyriakides, Campbell and Gagatsis 2000), the question of pedagogy has not been sufficiently explored in SER.

According to Thrupp (2001, p. 449) another drawback associated with the employment of multilevel modelling is the underestimation of effectiveness in schools with small numbers of students because of the way residual values get shrunk towards the mean (Thrupp, 2001, p. 449). This correction is made to statistically adjust for the relatively big measurement error associated with classrooms or schools serving a small number of pupils. However, teachers serving in these schools/ classrooms might feel disappointed

with this adjustment. *'While this shrinkage seems to be justifiable in statistical terms, for samples, if a teacher had obtained a very good result with a small group of students, he or she would certainly not want the results to be adjusted or downward'* (Fitz-Gibbon, 1996, p. 130, cited in Teddlie et al., 2000, p. 110).

Reynolds et al. (1996) summarise the criticisms of School Effectiveness studies in the United Kingdom as follows:

- The majority of British studies have collected data only within disadvantaged and deprived socio-economic contexts. They express the concern that 'the exclusion of very advantaged catchment-areas may have constrained variance in organisational practices at the school level, and might have also resulted in the generation of accounts of organisational functioning that are not necessarily applicable to all school types'.
- The absence of attempts to discern those classroom or instructional processes that might be related to outcomes may have affected the explanatory power of the models as well.
- The 'lack of interface between school effectiveness research and school improvement practice'
- The 'absence of more than rudimentary attempts at theory generation'

Reynolds (1992, p. 16) criticises SER carried out in Britain as follows: *'it has been unable to participate in 'cutting edge' debates about sensitivity to context that are a cen-*

tral feature of discourse in the field in the United States (Hallinger and Murphy, 1986; Wimpelberg et al., 1989)'.

MacDonald (in Elliot, 2000, p. 176) criticises international comparisons among pupils' educational outcomes, predominantly in mathematics and language, which employ a school effectiveness design '*as a means of justifying state intervention in the business of education*'.

But Goldstein and Woodhouse (2000) also criticised studies that do not conform with the requirements of a full multilevel design. Such a study was the 'Worlds Apart' report conducted by Reynolds et al. (1996), because among other things this study used a particularly small sample size and also it used a cross-sectional, instead of a longitudinal design. Goldstein and Woodhouse (2000) also state that SER should itself be very selective in endowing credit to other SER studies and findings because many SER studies and findings are derived through a less stringent methodological rigour.

***B) UNDERACHIEVEMENT OF ETHNIC MINORITY PUPILS AND PUPILS
FROM LOWER SOCIAL CLASS AND EFFECTIVE SUPPORT SCHEMES
AIMING TO ALLEVIATE IT.***

***1. 11. Describing underachievement of ethnic minority pupils and pupils from
lower social class***

'Equality of opportunity for all is a vital issue of social and economic importance to the whole of society... If any individual is denied the opportunity to fulfil their potential because of their racial, ethnic, class or gendered status it is now widely understood that society as a whole bears a social and economic cost by being deprived of the fruits of their enterprise, energy and imagination' (Gillborn et al., 2000, p. 6).

Equal educational opportunities imply that pupils participate at equal rates and achieve comparable educational outcomes with majority pupils, regardless of factors such as race, gender, family income or parental education.

In the UK the Swann report (Swann, 1985) found that ethnic minority pupils had lower performance than their British counterparts on almost every measure employed. As a consequence this finding rippled onto British educational policies, related to the issue of equal educational opportunities. Evidence drawn from six districts with large ethnic minority rolls showed clearly that *'West Indian children, on every measure, did very much less well in examinations than their white peers from the majority community'* (Swann, *op. cit.*). On the other hand, Asian children's attainment did not lag behind the attainment of their British counterparts.

The authors highlighted the point that the low attainment of West Indian children was partly due to the failure by educational authorities to identify and meet their educational needs (Swann, 1985, p. 5). It was argued that dealing with the issue of underachievement falls into the realm of the educational system:

‘In the long run, the Committee has no doubt that it is the educational system that has most to offer by way of help...It is surely realistic and right to expect that schools should address themselves to the task... Sensitivity to the needs of minority pupils is paramount’ (Swan, *op. cit.*, p. 12).

Furthermore, the committee stressed the need to shy away from single factor explanations, as multiple factors are at interplay, jointly shaping minority pupils’ attainment. Plausible explanations of lower attainment of ethnic minority pupils that were debated at the time include low intelligence, family structure, the materially and culturally disadvantaged West Indian home, racism (both in the society and the school), and the fact that districts do not address properly the issue of low attainment of ethnic minority pupils. In the report (p. 6) it is noted that much of the difference in IQ scores between children of West Indian origin and of British origin is related to differences between them in matters of socio-economic status. The report stresses that *‘when socio-economic status is taken into account, the extent of their underachievement in relation to their white peers is substantially reduced’* (Swann, *op. cit.*). In other words, both social class and ethnicity have a negative impact on the educational attainment of ethnic minority children. Sometimes lower social class and ethnicity interact, creating an additional negative effect for pupils from lower social class and minority ethnic background.

'What West Indians share with the white majority in lower social classes is a considerable measure of deprivation, but they suffer a substantial extra measure of deprivation, brought about by prejudice and discrimination' (Swann, 1985, p. 7).

The report concludes:

'The fundamental change needed is a recognition that the problem facing the educational system is not just how to educate the children of the ethnic minorities, but how to educate all children' (Swan, op. cit., p. 10).

In the USA the Connecticut State Board of Education in its five-year plan, 'Greater Expectations--Connecticut's Comprehensive Plan for Education 2001-2005', acknowledged that among the statutory goals for education the state should achieve resource equity and equality of opportunity, increase student achievement, reduce racial, ethnic and economic isolation, and improve effective instruction (Connecticut State Board of Education, Office of Board Matters, 2004).

The Connecticut Board acknowledges the need to further work in the domain of equal educational opportunities and adopt certain initiatives, such as: ensuring a high-quality teaching and administrative force; providing universal access to high-quality preschool; involving more parents and families in the education of their children; increasing the state share of revenues, particularly to these towns with students most in need; adopting educational monitoring initiatives and target setting for particular students' groups by districts and school, aiming to reduce achievement gaps between pupil groups.

Crawford (1997) summarising the Congressional Findings on the Bilingual Education Act, reveals that there are large and growing numbers of children and youth of limited-English proficiency in the USA, many of whom have a cultural heritage that differs from that of their English-proficient peers. The Bilingual Education Act that was created by the American Congress declares it to be the policy of the United States, in order to ensure equal educational opportunity for all children and youth and to promote educational excellence, to assist State and local educational agencies, institutions of higher education and community-based organizations to build their capacity to establish, implement, and sustain programmes of instruction for children and youth of limited English proficiency.

The purpose of this act was to educate limited English proficient children and youth to meet the same rigorous standards for academic performance expected of all children and youth, including meeting challenging State curriculum content and student performance standards in academic areas aiming to alleviate attainment disparities between minority pupils and majority pupils. This goal could be reached by:

- (1) developing systemic improvement and reform of educational programs serving limited English proficient students through the development and implementation of exemplary bilingual education programs and special alternative instruction programmes;
- (2) developing bilingual skills and multicultural understanding;
- (3) developing the English of such children and youth and, to the extent possible, the native language skills of such children and youth;
- (4) providing similar assistance to Native Americans with certain modifications relative to the unique status of Native American languages under Federal law;

(5) developing data collection and dissemination, research, materials development, and technical assistance which is focused on school improvement for limited English proficient students; and

(6) developing programmes which strengthen and improve the professional training of educational personnel who work with limited English-proficient students.

Van De Jong (1989) reported attainment differences for minority pupils in the Netherlands. He found that on average Surinamese, Turkish and Cape Verdian children scored 15 points, that is to say one standard deviation below the Dutch and Moroccan children scored 23 points below the Dutch.

Strand (1999 p. 180) refers to the literature on underachievement of ethnic minority children in Britain. According to Strand, the committee of inquiry into the education of minority children reported on the much lower examination attainment of African Caribbean school leavers in Britain. The committee concluded that:

‘West Indian children as a group are underachieving in our education system and this should be a matter of deep concern not only to all those involved in education but also to the whole community’ (Lord Swann, 1985).

Underachievement is often identified in the performance of ethnic minority pupils in relation to majority group performance. Assuming that ability is distributed in a similar way (that ability follows a normal distribution) among all ethnic groups enrolled in the schools of a given state, average attainment differences between majority and minority groups, especially after these ethnic groups have spent a considerable number of years

in the host country, reflect unequal educational opportunities provided to the minority in relation to the majority group.

Kim-Dong-il (2001), defines underachievement as the discrepancy between two standardized measures: intellectual capacity (e.g., IQ) and scores on standardized achievement tests. Broadly a pupil who is underachieving is not accomplishing at a level consistent with his/her abilities and as expected by reasonable and proper internal and external standards. According to Kim-Dong-il (2001), 'underachievement can then be defined as "performance, which does not measure up to the individual's level of aptitude" (Chaplin, p. 556), or performance below the "expected level"(American Heritage Dictionary of the English Language, p. 1395) indicated by that individual's performance on ability and aptitude tests'.

In England, Strand (1998, p 135), whose study adheres to the school effectiveness design found that ethnic minority pupils and pupils from lower social class (with free school meal entitlement) achieved lower scores at the end of Key Stage 2 than non disadvantaged pupils of the same reasoning ability. The study controlled for abstract reasoning score instead of initial attainment score and therefore differences in pupils' final attainment scores were adjusted for their ability. Many School Effectiveness studies carried out within the equal opportunities research tradition estimate the impact of initial attainment score along with other socio-economic characteristics on pupils' final attainment score; hence the finding of underachievement of disadvantaged groups is based on the fact that these groups make less progress in relation to their comparison groups, given that the initial educational attainment of disadvantaged groups is lower

than the attainment of their comparison groups (Mortimore et al. (1988), Bondi (1991), Hutchison (1993) and Strand (1997, 1999)).

Studies cited in Gillborn and Gipps (1996) proved that in the U.K. there are important differences in achievement between ethnic groups. GCSE results (for 1992 to 1994 inclusive) show Asian pupils performing less well on average than their white peers, but significantly better than the African-Caribbean group. Indian pupils are achieving levels of success consistently in excess of their white counterparts in some urban areas. Pakistani pupils are not achieving as highly as their white peers, etc.

The issue of underachievement related with ethnicity or social class has raised awareness in the English Department for Education and Employment. In 1997 the DfEE publication 'White paper for excellence in schools' stated that 'in order to overcome economic and social disadvantage and make equality of opportunity a reality, we must strive to eliminate, and never excuse, underachievement in the most deprived parts of the country. We must overcome the spiral disadvantage in which alienation from, or failure within the education system is passed on from one generation to the next' (DfEE, 1997, p. 3).

In the USA Eccels (1997, p. 68) associates the problem of ethnic minority underachievement with the problem of inadequate education and high dropout rates for ethnic minority pupils. This author states that the problem of underachievement goes across the board for many curriculum domains. She states that the latest round of National Assessment of Educational Progress (NAEP) has shown that Blacks and Hispanics by the age of 17 are testing four years behind their white peers in the sciences, although there

has been an increment in their relative position. Furthermore from the total Hispanic population across USA by age 13, close to 50% are already behind in school by at least one year.

The United States Department of Education (1994) based on the National Assessment of educational progress (NAEP) reported that gaps in the academic performance of black and white students appear as early as age 9 and persist through age 17.

However, there are different attainment patterns for different ethnic groups in the USA. According to the SAT scores, for the years 1976 to 1985 white students performed better than all other groups in terms of performance associated with verbal ability, while Asian and White students performed better than all other groups in mathematics.

Olnec (1995) reviewed several studies carried out in the USA. Caplan et al. (1991) found that children from Asian countries (the boat people) despite having been in the USA only an average of 3.5 years and usually from households in which no one came to the USA knowing English, received grades averaging slightly over B, and scored at the 54th percentile on the California Aptitude Test (CAT). On the mathematics section of CAT they scored at the 72nd percentile.

Portes and Rumbaut (1990) (in Olnec, 1995, p. 323) performed comparisons between students from diverse ethnic groups in the USA. The data referred to 39000 high school students in San Diego during 1986-1987. This data revealed a familiar pattern of lower grade-point averages (GPA) among African Americans, Pacific Islanders, and Mexican – origin students, whereas higher grade point averages (GPA) were found among white Anglos and highest GPAs among Korean, Chinese and Asian Indian students.

1. 12. Scope of support education

Among the list of characteristics of unusually effective schools provided by Levine (1990, p. 579) appear characteristics related to the availability of support instruction at the school level, namely the availability and effective utilisation of instructional support personnel, the creation of extra time for reading, language and mathematics, the emphasis on multicultural education and sensitivity and finally rigorous and equitable student promotion policies and practices.

Support education aims to raise the educational outcomes of foreign/repatriated pupils so that these reach parity with the outcomes of their native counterparts, or raise the educational outcomes of native and foreign pupils with accumulated learning gaps. This is an action of positive discrimination fighting social exclusion and poverty. Hargreaves (1982) and Shaw (1984) (both cited in Deem, 1986) stressed that most schools focus only on developing the potential of individual pupils and thus fail to address group inequalities such as class and gender. *'From the work of Floud and Hasley in the 1950s onwards (Halsey et al., 1961) the dominant concern has been with working class failure, so-called social exclusion'* (Whitty, 2001). Whitty (*op. cit.*) considered affirmative action as necessary in order to tackle the extent to which working class children continue to be denied the opportunities open to middle class children, and so as to enable working class children to succeed.

Cotton (1995) explicitly refers to promoting equity as an effectiveness enhancing condition. Scheerens and Bosker (1997, p. 180) comment that Cotton by emphasising equity

is giving prominence to an equitable distribution of achievement for different groups of pupils.

Murphy (1989, p. 33) argues that equity issues should be separately addressed within an educational reform. In addition, the author argues for the necessity that in the case of the American school system the positive effects of reform should flow disproportionately to disadvantaged students.

The European councils of Lisbon (2000) and Feira (2000) declared the fight against poverty to be one of the central elements of the European social model (Camilieri, 2003). In this framework, the percentage of persons with only compulsory or with less than compulsory education in each European member state has been set as an indicator of social inclusion (Stanton, 2003). And since poorer educational outcomes are found among pupil groups with higher dropout rates (Osborne, 2001), raising the attainment of disadvantaged groups would raise the social inclusion indices for these groups.

In England in addition to Swann (1985), Mortimore et al. (1988), Bondi (1991), Hutchison (1993), Strand (1997, 1999), Gillborn and Gipps (1996), and in the USA Portes and Rumbaut (1990) and Olneec (1995) have provided evidence of under-achievement for minority groups demonstrated in substantial attainment and/or progress differences between minority and majority groups.

In Greece, the current study is the first primary school effectiveness study, carried out within the equal opportunities research tradition.

Furthermore, it is morally and ethically unjustifiable to deprive from support instruction pupils (from both majority and minority groups) who due to poverty, poor health, or family problems (associated with belonging to a lower social class), have fallen behind in their attainment compared with the rest of the class. *'Children living in poverty have less access to formal educational opportunities, fewer resources, greater health problems and developmental delays, all of which negatively impact on educational outcomes'* (Ford, 1996, in Borman et al., 1998, p. 11).

In Greece *'the highest levels of poverty appear in households where the head has not completed primary education or s/he has only been educated up to the primary school level, in relation to households with heads who have completed higher educational levels'* (Barloudas et al., in Tressou, 1998, p. 666).

According to Barloudas, (*op. cit.*) in Greece the percentage of poverty for graduates of primary education is 51.6%, whereas for the graduates of secondary education first grade is 8.3% and for the graduates of secondary education, second grade is 8.6%.

Ford (1996, in Borman et al., 1998, p. 11) refers to Montgomery and Rosi (1994) who claim that resources in one system (at home or in the school) may mediate risk factors in the other. In other words an intellectually stimulating home can compensate for inadequate schooling, while a supportive school may attenuate the effects of social disadvantage.

Gilborn and Gipps (1996) identified social class as a factor directly related to academic achievement: *'the higher the social class, the higher the achievement'* (Gilborn and Gipps, 1996, p. 16). The effect of social class on attainment can be detected even when

the effects of gender and ethnicity are accounted for. *'Social class is strongly associated with achievement regardless of gender and ethnic background'* (Gilborn and Gipps, 1996). Furthermore, pupils' gender is associated with attainment differences; boys, on average, have lower attainment than girls. In the UK Gillborn and Mirza (2000, p. 22), refer to the finding that boys in UK perform at a significantly lower level than girls as *'the new gender gap'*.

Social class affects the attainment of every child very early on, in kindergarten.

'The indicators of inequality in school readiness show that children start kindergarten with different levels of reading and mathematical skills, often related to gender, race/ethnicity, socio-economic status and age' (Coley, 2002, p. 61).

Social class, ethnicity and gender effects persist and they are detectable when the pupil leaves primary school (at the end of Key Stage 2 in UK). These effects persist throughout primary and secondary education, as shown by the Swann Report (Swann, 1985), Gillborn and Gipps (1996), as well as by several School Effectiveness studies such as by Mortimore et al. (1988) in England. The Junior School Project (Mortimore et al., 1988) found that overall, among the variables age, gender, ethnicity and social class, it was social class that accounted for the main difference between groups of pupils.

If minority pupils are not assisted in acquiring early on the language that is used as a medium of instruction, linguistic barriers that inhibit their understanding remain for a long time and these pupils do not have full access to the curriculum taught in the mainstream class.

According to Ellis (1990) second language learners acquire the second language through formal instruction and informal exposure to the second language. *'Learners require both formal instruction and informal exposure and that the two together work better than either on its own'* Ellis (1990, p. 32).

Ellis (1994, p. 653) has reviewed the results of many other studies (for example Harley, 1989; White et al., 1991), which have shown that formal instruction improves grammatical accuracy of second language learners, and that the gains that learners make can be durable. Grammatical accuracy is related to pupils' performance in tests.

In the Greek setting, minority pupils learn the Greek language through formal exposure when they acquire new vocabulary and structures of the Greek language by attending lessons delivered in the mainstream class or by attending lessons delivered in coach classes and reception classes where support teachers strengthen concepts taught in the mainstream. On the other hand they learn Greek through formal instruction in cases where support class teachers teach Greek as a second language to them in coach classes and in reception classes;

In addition, by not assisting low achieving pupils to cover their learning gaps in core subject, like language and in mathematics, cognitive barriers that inhibit their understanding in subjects taught in the mainstream remain.

By not providing support teaching to native low achieving native pupils and to low achieving foreign pupils who have already acquired an adequate command of the language used as a medium of instruction, principals turn a blind eye to the effects of social class. The effects of social class are interlinked with the effects of ethnicity and the

effects of gender (Gillborn and Gipps, 1996). *‘If schools do not provide the necessary language services to break down the linguistic barriers, social inequalities will continue’* (Tollefson 1991). Support teaching can be considered as a mechanism aiming to alleviate educational and social disadvantage, thus redistributing educational outcomes.

To establish a policy focusing on equality of opportunity, support provision should target all pupils with low attainment. Among the pupils in need of support, other than pupils with statements of special educational needs⁵, three groups with different needs can be distinguished: foreign and repatriated pupils with limited competence in their second language; minority pupils who might have acquired the language that is used as a medium of instruction but over the years may have accumulated learning deficits; and majority pupils with accumulated learning gaps. Support teaching should target all pupils in need of support, in a comprehensive framework of intervention. The intervention implemented in each school should target each pupil with low attainment according to the origin of the pupil’s needs. Targeting the needs of all pupils does not imply that all pupils should be placed together in the same setting, in order to receive support instruction. For example, if the school wishes to target the needs of beginner second language learners and advanced second language learners in order to raise their attainment, then support teaching may be organised differently and delivered through distinct approaches, in line with the recommendations of language pedagogy.

Many of the countries that realised that minority pupils were experiencing difficulties in their educational systems started to provide additional programmes and resources tailored to the needs of these groups. In the Netherlands the Cultural minority policy and the Educational Priority Policy were initiated. The aim of the latter policy was to reduce or eliminate the educational disadvantage of pupils that is a consequence of social, economic and cultural circumstances by raising the Dutch language and arithmetic levels in primary schools, since several studies made it clear that these were the main inhibiting factors with regard to disadvantaged children's school careers. The main target groups were non-indigenous and indigenous children whose parents have a low educational and occupational level (Driessen and Mulder, 1999, p. 38).

According to these authors in Australia important programmes include the anti-poverty Disadvantaged Schools Programme (DSP), which aims to stimulate literacy and numeracy, enhance life experience and self-confidence and improve interaction between schools and the neighbourhood, and the National Equity Programme for schools, which aims at creating opportunities for disadvantaged children.

In West European countries various initiatives have been undertaken to improve the educational level of disadvantaged children. Since 1981, in France for instance, so-called ZEPs (Zones d' Education Prioritaires) –Educational Priority Areas – have been

5 According to Greek Law 2817/2000 (FEK 78/14 / 03 / 2000) pupils who are suspected of having special educational needs are sent to special pedagogical centres belonging to the municipality, or to the newly established 'centres of diagnosis, assessment and certification' (KΔAI) in order to be thoroughly assessed. These centres after administering to the child a diagnostic assessment tool, they compose a 'statement of special educational needs'. With this statement the pupil can enroll in the 'sections of integration' operating in his/her mainstream school, which have replaced special classes. Pupils without a statement of special educational needs enroll in the above sections with the consent of the majority of the teachers serving in the school unit. The above sections operate for a limited number of hours, and for the rest of the school day pupils with special educational needs enroll in their mainstream class.

in operation. In ZEPs professionals and volunteers from such different areas as teaching, social services, health care and the police force work together. Activities include reading programmes, promoting citizenship, and extra – school help. Since 1991 Belgium has had an Educational Priority Policy aimed at improving the opportunities of ethnic minorities in nursery, primary and lower secondary schools. Schools with an approved action plan, and at least 20 pupils of the target group, are entitled to extra teaching staff.

The English OFSTED (2000) described effective support as one that involves close oversight of the pupils' academic progress, based on the analysis of pupils' needs, making good use of data and engaging parents effectively. Pertinent additional teaching and generous extra-curricular activities should back up a support teaching initiative.

In England the School Curriculum and Assessment Authority (SCAA) acknowledged that: *'Effective teaching of English⁶ draws on its purposeful use across the curriculum; and Effective use of English is vital to the teaching and learning of every subject'* (SCAA, 1996, p. 2).

Support education aims to provide pupils with extended instructional time opportunities and more individualised instruction or more instruction adapted to their actual level and needs. Support instruction aims among other things⁷ to alleviate the impact of poverty (which is associated with pupils having a poor socio-economic background), and/or the impact of ethnicity on pupils' educational attainment. Usually pupils belonging to foreign ethnic groups have poor skills in the second language.

⁶ The main language of instruction

According to Reynolds (2000, p. 250) countries situated in the Pacific Rim such as Korea, Japan and Taiwan that scored well above average in International School Effectiveness studies such as in the Second International Mathematics Study (Scheerens et al., 1989) operate with the prevalent belief that *'all children are able to acquire certain core skills in core subjects and that there is no need for a trailing edge of low performing pupils'*. This belief is contrasted with the commonly held view in western societies that since pupils' performance follows a normal distribution, or since pupils' ability is normally distributed, a certain percentage of pupils are expected to attain at a below-average level. However, although this assumption is statistically correct (50 per cent of the pupils tested would attain below the 50 percentile), support teaching can ensure that the variance in pupils' attainment distribution is limited; consequently attainment disparities between average achievers and low achievers can be reduced with the provision of adequate support teaching to low achievers. If this happens, there would be no trailing edge in the lower part of all pupils' attainment distribution.

The Australian State provides support teaching throughout primary school:

'Literacy development remains a priority for all students as they progress through the grades' (NSW, 1997, in Rowe, 2000, p. 9).

The official publication of English Department for Education and Employment, *'Excellence in schools'* (DfEE, 1997, p. 22), points out that *'sharper focus on literacy and numeracy in the curriculum becomes a priority for action'*. The learning of most other

⁷ Compensatory education can also aim to reduce gender effects, or effects associated with the child having special educational needs.

subjects taught in the mainstream class depends upon having acquired adequate skills in literacy and numeracy.

1. 13. A summary of support teaching provision in Greece

Many foreign/repatriated pupils who are enrolled for the first time in a Greek school have limited competence in Greek (their second language), a condition related to their low attainment in other subjects as well. Underachieving pupils may also be foreign/repatriated pupils who, after having spent some years in a Greek primary school, have learnt the language, but over the years, have accumulated learning gaps in basic subjects.

Low achieving pupils may be either Greek or foreign pupils who did not make the right start in school because of an initial disadvantage in their educational attainment, or who have accumulated learning deficits acquired throughout their schooling. Davie et al. (1972), Strand (1999), Cowie and Croxford (1999), and Meijnen et al. (2003) found that attainment differences have already appeared at entry into primary school. Therefore if these pupils are not assisted in order to alleviate these attainment differences, because often disadvantaged groups make less progress in comparison to non-disadvantaged majority groups as it was shown in the study of Strand, (1999) attainment differences are bound to increase even more in the future, as pupils' initial attainment score is the strongest predictor of their final attainment score (Goldstein, 1997).

As mainstream class teachers usually adjust the level of instruction to the level of the average pupil, foreign/repatriated, and low achieving Greek pupils may find the mainstream class lessons too advanced for them to understand.

Hence, foreign/repatriated pupils need to be taught the Greek language in an intensive manner so that they can understand the curriculum taught in the mainstream class. Low achieving Greek and foreign pupils should be assisted in covering their learning gaps in major areas of the curriculum. Foreign/repatriated pupils often have accumulated learning gaps from the period they did not fully comprehend the Greek language.

In cases where foreign/repatriated pupils have limited competence in the second language, they are targeted as a separate group. In the support class, the main focus of instruction is to teach them Greek as a second language, or to provide them with support to understand the lesson in the mainstream and complete their homework and assignments. Missing language structures and vocabulary, or missing concepts, may inhibit them from learning in the mainstream class setting. As foreign/repatriated pupils learn Greek (the language of the school), they become more able to acquire content and language structures from the mainstream class setting.

There are distinct initiatives that may be undertaken to address the needs of low achieving groups, in order to reduce the attainment gap between low and high achieving pupils in Greek primary schools. In 1983, reception classes and coach classes⁸ were institutionalised in law: (N1404/83 no 45). The rationale behind the creation of reception classes and coach classes in primary schools was *‘to enable foreign and repatriated pu-*

⁸ In Greek coach classes are called *‘frontistiriaka’* sections.

pils to be smoothly integrated into the Greek educational system' (Damanakis, 2001, p. 61) but the main goal was to assist these pupils in learning the Greek language. Reception classes function in parallel with the mainstream, whereas coach classes operate after the end of the mainstream class programme (2 pm). Support teachers working in reception classes and in coach classes usually identify gaps in pupils' understanding, repeat and elaborate concepts taught in the mainstream class, while some support class teachers give a high priority to the teaching of Greek to foreign/repatriated pupils. In this way these pupils become more able to make sense of and appropriate further knowledge in the mainstream.

However, there are initiatives that schools can undertake to address the educational needs of all pupils, high and low achievers, with an emphasis placed on the needs of children with working parents. While the introduction of support schemes for low achieving pupils is justifiable as it addresses equity concerns, it does not preclude the introduction of enrichment activities for all the pupils, such as the activities included in the recently introduced whole-day school in Greece. The later initiative is established in Law 2525/97 in ΦΕΚ 188/2Α article 4. Its functioning is further specified through ministerial decree Φ13.1/ 767/ Γ1/884/3-9-1998 and ministerial circulars Φ13.1/717/ Γ1/742/21-9-1999 and Φ13.1/897/ Γ1/694/6-9-2000.

Under the *blanket term 'whole day school'* reference is made to *two separate initiatives*: The first refers to schools with compulsory after school attendance and schools adopting it are officially referred to as 'whole day schools'. The second initiative refers

to schools of extended timetable in which after school attendance is optional for the pupils.

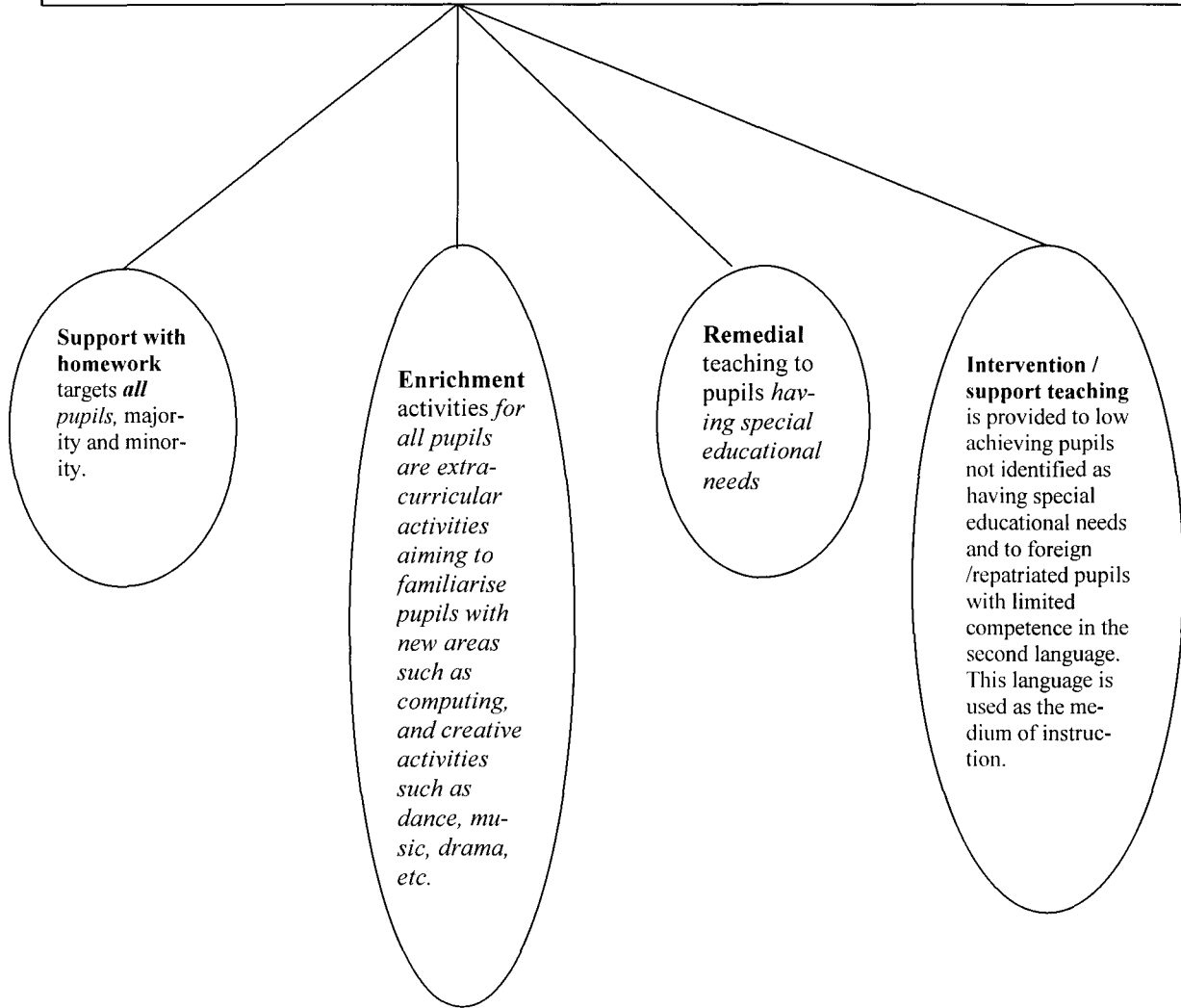
The first initiative requires that pupils attend after-school programmes in their school on a *compulsory basis* (Law 2525/97 in ΦΕΚ 188 Α) until 4 p.m. These schools offer to the pupils an additional programme of nine hours per week.

The second initiative refers to enrichment programmes and to support with homework sections organised within the framework of *whole day school*⁹. Additional support with homework programmes aims to support all pupils in all curriculum areas not mastered during the mainstream class lesson. Enrichment activities do not have a pre-defined content, but according to Stamelos (2002, p. 82) these can be based on fifteen subjects.

These activities can have a cultural/artistic character, such as musical, kinaesthetic and theatrical education, or they may be athletic games, events, activities and dances, IT literacy (computing), health education, road safety, mythology, foreign languages, etc. Through these activities foreign/repatriated pupils may be assisted in further acquiring the Greek language through their participation and use of Greek in more informal settings. The involvement of foreign/repatriated pupils in activities that require handling or manipulation of materials facilitates the acquisition of their second language.

⁹ See definition in glossary

Support teaching
(provided in addition / extension to or restructuring the mainstream class)



1. 14. Support education programmes in other countries

1. 14. 1. Title I

Title I is a major American support teaching programme initiated in 1964. It aims to improve education for pupils at risk of school failure in low-income communities. Categorical aid is employed to provide extra assistance for disadvantaged children—migrant children, children for whom English is a second language, delinquent and neglected children, and children with mental and physical handicaps. Typically title I offers low achieving pupils approximately 30 minutes of support instruction per day (Borman et al., 2001, p. 51).

The New American Legislation, '*No-child-left-behind Act*' (United States of America - Congress House, Boehner, 2001), has revised title I services. This Legislation employs accountability systems¹⁰ and initiates high-quality academic assessments measuring attainment and progress in order to alleviate underachievement and improve the academic attainment of disadvantaged groups. Among other groups, Limited English-Proficient children, migrant children, children with disabilities, young children in need of reading assistance and low performing children in general are mentioned in this Act.

¹⁰ Educational accountability aimed at closer monitoring of the educational outcomes produced by schools.

More specifically, among other things, Title I programme aims to:

- Meet the educational needs of low achieving children in the USA's highest level of poverty schools, limited English-proficient children, migrant children, children with disabilities, Indian children, neglected and delinquent children, and young children in need of reading assistance;
- Close the achievement gap between high and low-performing children, especially the achievement gaps between minority and non-minority students, and between disadvantaged children and their more advantaged peers;
- Hold schools, local educational agencies, and States accountable for improving the academic achievement of all students, and identifying and turning around low-performing schools that have failed to provide a high-quality education to their students, while providing alternatives to students in such schools to enable the students to receive a high-quality education;
- Distribute and target resources sufficiently to make a difference to local educational agencies and schools where needs are greatest;
- Improve and strengthen accountability, teaching, and learning by using State assessment systems designed to ensure that students are meeting challenging State academic achievement and content standards and increasing achievement overall, especially for the disadvantaged;

- Provide children an enriched and accelerated educational program, including the use of school-wide programmes and additional services that increase the amount and quality of instructional time;
- Significantly elevate the quality of instruction by providing staff in participating schools with substantial opportunities for professional development;

(United States, Congress House, Boehner, 2001, p. 17)

It is interesting how the revised Title I programme combines accountability, school improvement and additional services for the disadvantaged, Limited English-Proficient pupils and low achieving pupils enrolled in high level of poverty schools, coupled with up-to-date assessment and good teaching, to ensure that all pupils have a fair, and equal chance to obtain high-quality education. The requirements for Limited English-Proficient pupils are demonstrated further in the above-mentioned act (on p. 284), where it is written that each State Educational Agency shall develop annual measurable achievement objectives for Limited English-Proficient children that relate to such children's development and attainment of English proficiency, while meeting student academic achievement standards. More specifically each state educational agency will publish the annual increases in the number or percentage of pupils making progress in English; in the number or percentage of pupils attaining English proficiency at the end of the school-year, and the percentage of Limited English Proficient pupils who are making adequate yearly progress. Additional state assessments and standards are required, for which the states will have to undertake the costs; among others, the development of

assessments of English language proficiency for Limited English-Proficient pupils, as well as to ensure continued reliability and validity of current state assessments (p. 457).

In the case of Limited English-Proficient pupils the Act proposes the following in Sec. 3101:

- To help to ensure that children who are Limited English Proficient, including immigrant children and youth, attain English proficiency, develop high levels of academic attainment in English, and meet the same challenging State academic content and student academic achievement standards as all children are expected to meet;
- To assist all Limited English-Proficient children, to achieve at high levels in the core academic subjects so that those children can meet the same challenging state academic achievement standards as all children are expected to meet;
- To develop high-quality language instruction educational programmes designed to assist state educational agencies, local educational agencies and schools in teaching Limited English-Proficient children and serving immigrant children and youth;
- To assist state educational agencies and local educational agencies to develop and enhance their capacity to provide high-quality instructional programmes designed to prepare Limited English-Proficient children, including immigrant children and youth, to enter all English instruction settings;
- To assist state educational agencies, local educational agencies and schools to build up their capacity to establish, implement, and sustain language instruction educa-

tional programmes and programmes of English language development for Limited English-Proficient children;

The 1988 Hawkins-Stafford Amendments attempted to establish students' grade-level proficiency in both 'basic' and 'more advanced' skills (U.S. Department of Education, 1996). These amendments encouraged districts to enhance student services and advocated frequent and regular coordination between Title I and regular education staff.

Among the strategies adopted in the renewed Title I programme are the extended day, extended year, and summer programmes

(<http://www.ed.gov/pubs/NatAssess/exsum2.html>, 1996 and Hamby, 1989).

All the above programmes increase the effective teaching time, or the instructional time tailored to disadvantaged pupil groups. An accelerated, high quality curriculum and increased flexibility in decision making at the school level are other measures put forward, so that teachers become professionals who choose the most promising practices and initiatives for the benefit of their pupils.

Borman and Agostino's (2001, p. 47) findings suggest that although Title I programme has not achieved its promise to eradicate differences between the disadvantaged and the more advantaged groups, it has succeeded in partly alleviating the educational and social disadvantages.

The effect size¹¹ of Title I programme has been estimated to have been nearly 0.15 during the early 1980s. The effect size is an index widely employed in assessing the effectiveness of educational and in many cases of experimental programmes. It is usually estimated during the evaluation of a programme. It expresses the proportion of a standard deviation separating the experimental from the control group in terms of the educational outcome at stake. It is used as a core criterion in the evaluation of a support - teaching programme; more specifically it addresses the question whether attendance in the programme has improved pupils' educational attainment.

11 The effect size is a particularly useful tool in assessing the effectiveness of experimental programmes and expresses the proportion of a standard deviation separating the experimental from the control group in terms of the educational outcome at stake (in many cases the educational outcome is attainment).

The effect size of a given educational programme is estimated according to the following formula:

$$ES = \frac{(\text{progress of experimental group}) - (\text{progress of control group})}{SD(\text{progress of control group})}$$

1. 14. 2. Success for All

‘Success for All’ is a programme designed by Slavin and his colleagues in Baltimore, USA ‘*to ensure that every child who enters school, regardless of home background, will succeed in basic skills in the early grades and then maintain that success through the elementary years*’ (Slavin & Leighton, 1990, p. 1).

According to Slavin and Leighton (1990), ‘Success for All’ targets pupils in the early grades as other remedial programmes do. It usually concentrates on pupils enrolled in kindergarten and in 1st, 2nd, and 3rd grades of primary school. ‘*Success for All focuses on prevention, early intervention and long-term professional development, instead of remediation*’ (Slavin and Madden, 1998). In other words, the programme puts emphasis on early prevention and on timely intervention so that low achieving pupils catch up with the mainstream curriculum, without having to enrol in a special education programme. The programme includes the following elements: one to one tutoring from certified teachers for students who have difficulties in reading, frequent assessment of all pupils (every eight weeks), family support services, grouping and regrouping the mainstream class pupils by level for reading, cutting down the class size in the mainstream and other interventions.

The ‘Success for All’ programme and its mathematics sub-component ‘Roots and Wings’ (Slavin, R., N. Madden, et al. (1996), Slavin and Madden (1998), Slavin, and Fashola (1998)’ focus on the development of literacy and numeracy skills. The ‘*Success for All*’ programme aims to raise standards in low achieving schools but also to raise the attainment of disadvantaged groups of pupils. ‘Success for All’ simultaneously targets

every pupil in the mainstream class and low achieving pupils as a separate group by providing them with individual tutoring during 20-minute sessions that do not coincide with the reading or mathematics instruction in the mainstream. Pupils in the mainstream class are regrouped for reading instruction so that instruction is conducted in groups of no more than 15 pupils and at one reading level at a time. Regular assessment of pupils takes place (usually every eight weeks) so that pupil groups as homogeneous as possible are formed.

This programme has shown evidence of good progress for pupils enrolled in schools that implement this programme, but even more progress has been reported for disadvantaged groups of pupils. This evidence of good progress is reflected in the programme's effect size.

Effect sizes associated with the 'Success for All' programme on average amount to about one half of a standard deviation in relation to equivalent schools not having adopted the 'Success for All' programme, at all grade levels. 'Effect sizes for students in the lowest 25% of the grades were particularly positive, ranging from $ES = +1.03$ in the first grade to $ES = +1.68$ in the fourth grade' (Slavin et al., *op. cit.*, p. 56).

Low achieving pupils enrolled in the 'Success for All' programme are pulled out of the mainstream for individual tutoring, but care has been taken so that pupils do not miss out on literacy or numeracy in the mainstream. According to Slavin and Madden (1998) '*individual pupils' tutoring occurs in 20-minute sessions during times that do not coincide with the reading or mathematics lesson in the mainstream*'.

The 'Success for All' programme resulted in reduced retention rates and special education referrals in schools primarily serving disadvantaged African American students (Slavin, Madden, Karweit, Dolan & Wasik, 1990; Slavin, Madden, Karweit, Livermon & Dolan, in press referred to in Slavin and Leighton, 1990, p. 1).

Although initially 'Success for All' was used with majority pupils from poor families, at a later stage, it was used for minority pupils who were Limited English Proficient as well, with similarly satisfying results (Slavin and Madden, 1998). Limited English Proficient pupils were taught English as a second language receiving one-to-one tuition or tuition in small groups, based on mainstream curriculum materials. 'Success for All' employs experienced tutors to support pupils' success in reading, as tutoring is a well-known form of intervention for low achieving pupils. Tutoring occurs in 20-minute sessions during times not coinciding with the teaching of reading or mathematics in the mainstream class.

In general, tutors support pupils on the same objectives taught in the mainstream classroom. Tutors also seek to identify learning problems and use different strategies to teach the same skills.

According to Slavin and Leighton (1990), daily and during two-hour reading/ language arts periods, tutors serve as additional reading teachers to reduce class size for reading. Reading teachers and tutors use brief forms to communicate about pupils' specific problems and needs and meet at regular times to co-ordinate their approaches with individual children. Initial decisions about reading group placement and the need for tutoring are based on informal reading inventories that the tutors give to each pupil. First

graders receive first priority for tutoring, on the assumption that the primary function of the tutor is to help all the pupils to be successful in reading first time, before they become remedial readers.

All pupils enrolled in 1st, 2nd and 3rd grades are regrouped for reading. Pupils who are assigned to heterogeneous, age-grouped classes with class-sizes of about 30 most of the day, during the regular two-hour reading/language arts period are regrouped according to reading performance levels into reading classes of 15 pupils all at the same level.

‘Regrouping allows teachers to conduct the lesson at one reading level’ (Slavin and Leighton, 1990, p. 2). This maximises the effective teaching time, as pupils are able to understand the content of instruction most of the time. Slavin and Leighton (1990, p. 2) state that all the pupils are regularly assessed and regrouped in such a way that the groups created are as homogeneous as possible. Minority pupils who have limited competence in English are given support related to the mainstream curriculum. Regrouping increases the time of direct instruction. A programme facilitator worked in each school to facilitate the operation of the programme. Facilitators visited mainstream classes and tutoring sessions frequently to provide solutions to teachers’ and tutors’ individual problems. Teachers and tutors received regular training on such topics as classroom management, instructional pace, and implementation of the curriculum (Slavin and Madden, 1987). Tutors received manuals, which provided them with explicit instruction of how to go about the intervention programme.

1. 14. 3. Reading Recovery

Marie Clay was the first developer of Reading Recovery in New Zealand in 1985 (Wasik and Slavin, 1993, p. 181). Reading recovery provides one-to-one tutoring to first and second graders who score in the lowest 20% of their classes on a programme-developed diagnostic survey. The programme has also consistently shown benefits for pupils learning English as an Additional Language (EAL) in general (Cline and Shamsi, 2000, p. 58). Tutors are certified teachers who receive training for two and half hours per week for an entire academic year. Pupils are tutored for 30 minutes per day until one of two things happens: If pupils reach the level of performance of their classmates they are discontinued. If they receive 60 lessons without achieving this level of performance, the pupils are released from the programme but considered 'not discontinued'.

Proponents of Reading recovery such as Swartz and Klein (1996) suggested that intervention efforts should concentrate on the first grades, since intervention lasts for a short period of time while participation in the programme at this stage has a lasting effect on pupils' attainment. Kennedy, Birman, and Demaline (1986 in Crevola and Hill, 1998, p. 134) suggested that efforts to correct such problems beyond the third grade are largely unsuccessful.

The Reading Recovery model includes the following reading components: perceptual analysis, knowledge of print conventions, decoding, oral language proficiency, prior knowledge, inference making, reading strategies, meta-cognition and error detection, and error correction strategies (Clay, 1979, cited in Wasik and Slavin, 1993, p.193).

Arlington (1988) has identified the following features of Reading Recovery: Early intervention, Short-Term Extra Help, Building on strengths, Independence, Flexibility and Responsiveness, Action-Oriented, Enabled Participation, and Accelerated Progress (King, 1991; Gay Su Pinnell, 1988; Brophy and Alleman, 1990).

Pinnell (2000) described ten principles included in Reading Recovery and in literacy programmes in general:

Phonological awareness can be described as the child developing the ability to hear the sound in words. *Visual Perception of letters* refers to the pupils' ability to perceive and identify letters of the alphabet, and to learn how to look at print. *Word recognition* refers to building a small repertoire of words known in detail by the pupil, which the pupil can recognise quickly. *Phonics/ decoding skills* refer to pupils learning and solving complex letter-sound relationships in several different ways, and being taught to apply knowledge in reading and writing. *Phonics/structural analysis* refers to pupils using structural analysis of words and learning spelling patterns. *Fluency-automaticity* in pupils' reading and writing develops after a certain time passes, as the learners have to gradually speed up. *Constructing meaning from print* is a technique that has to be employed from the beginning, so that pupils can comprehend what they read.

Reading recovery presents a balanced structured approach to early literacy development so that literacy develops along a broad front and pupils can apply skills in reading and writing.

Pinnell et al. (1988), pp. 10-11 describe the tutoring process in the following manner: Initially, the child reads and rereads easy and familiar books. The materials are story-

books with natural language rather than controlled vocabulary and gradually proceed to more advanced materials. The teacher keeps a running record of a child's performance in oral reading; then running records are examined closely, analysing errors and paying particular attention to behaviour such as self – correction. In this way, s/he determines the strategies the child is using to gain meaning from a text. This assessment provides an ongoing picture of the progress a child makes. An accuracy check tells the teacher whether the text was well selected and introduced the day before.

A high proportion of successfully discontinued children continues to make progress for at least two full years after the individual reading recovery intervention has taken place (Pinnell et al., 1991, *p.* 37).

1. 14. 4. Levine's approach

Levine first worked in the 1960's in the UK, after the first waves of immigrants settled down there. Levine (1990) firmly believed that minority pupils should always be assisted in the mainstream class, along with native pupils for several reasons. Minority pupils should be exposed to the same curriculum as native pupils, interaction opportunities between foreign and native pupils should be provided, so that the former socialise and acquire friends from the majority group.

Levine was against teaching minority pupils in a separate setting in 'Special English' classes, away from the mainstream as activities and tasks taught there were too simplified in comparison to the requirements of the mainstream class. In contrast, she believed that it is more productive for both language and other lessons *'to derive the language syllabus from the content and activities of mainstream curriculum lessons, rather than look for content and activities with which to clothe any pre-constructed language syllabus'* (Levine, 1990, p. 22).

Levine (1983) suggested that second language learning should take place in the mainstream where teachers arrange for minority pupils to learn and practice the second language in communication and interaction with native speakers as they engage with the subject matter, in real meaningful contexts. Levine (1990) condemned special English or classes which withdrew foreign pupils from the mainstream as *'functioning not to help the minority pupils in mainstream'*, but to *'help the mainstream by withdrawing difficulties from it'* (p. 18). She believed that support teachers working in the main-

stream class needed to *'provide foreign pupils with new language, as the pupils needed it in comprehensive contexts'* (p. 21). Hence, foreign/repatriated pupils with limited competence in the second language should be assisted in completing the tasks of the mainstream class while they acquire the vocabulary and the language structures that would enable them fully to understand the curriculum in the mainstream class.

Levine was a proponent of a rich learning environment in which pupils could learn to use their second language (English) through employing it for academic purposes.

Providing assistance in the mainstream would enable minority pupils to understand the mainstream curriculum. In the mainstream, pupils *'might have even wider, more productive access to a range of language and a range of opportunity to develop a repertoire of skills than in the special language situation where everyone was in the process of learning to use English (the second language)'* (Levine, *op. cit.*, p. 22).

Carrying out the support instruction in the mainstream may guarantee that minority pupils are taught according to the mainstream curriculum, so that the lessons taught in the support and in the mainstream class reinforce each other. It may increase the support teacher's awareness of the mainstream class teacher's objectives, so that s/he ends up teaching the same goals. However, there is no guarantee that providing assistance to minority pupils within the mainstream class setting will invite the support class teacher to teach the same concepts taught by the mainstream class teacher, thus enhancing understanding of the content taught within the mainstream class. Actually, *'in the absence of good collaboration, the support teacher 'may simply withdraw the child and work independently, or two teachers may be working in a room, but in effect operating*

independently, or two teachers may be working in a room, but in effect operating withdrawal within the classroom' (Ireson, 1992, p. 57).

In practice, to support this arrangement a large number of support teachers may be required, which might prove hard to find; this is evident in many British schools where aides¹² are employed as support assistants. The employment of aides instead of teachers may compromise the quality of instruction provided in the support setting.

In addition, budgetary constraints would make it difficult for districts to provide support instruction for every mainstream class having minority pupils in need of support, especially in mainstream classes with relatively few such pupils.

Even in classrooms with many minority pupils with limited competence in the majority language it might prove difficult to sort the different pupils into homogeneous groups so as to maximise effective (support) teaching and thus effective learning time. If foreign /repatriated pupils from many grades are grouped together according to their current level of competence in Greek, then relatively homogeneous groups can be formed and thus many pupils at the same level of language competence may profit from the services of one teacher, thus providing more opportunities to fine-tune the lesson in order to meet pupils' needs.

12 According to Fox (1998, p. 28) 'personal qualities rather than qualifications and experience, except perhaps for grade C in GCSE English are seen as the main criteria for entry to the profession' for learning support assistants'.

1. 14. 5. Programmes supporting Numeracy

Cockroft (1982, par. 398, referred to in O'Donoghue, 2002, p. 47) defined Numeracy as 'the mirror image of literacy'. Literacy and Numeracy were seen as personal attributes that were needed to support the life-long aspirations of an educated person and that included communication between the so-called 'two cultures'. Literacy and Numeracy were seen as overlapping complementary attributes.

MathWings is a numeracy programme developed by Robert Slavin and his associates as a part of the 'Success for All' programme. MathWings uses cooperative learning at all age levels while incorporating problem solving in real situations, skill practice and reinforcement, calculator use, alternative assessments, writing, connections to the literature and other disciplines and application to the students' world and personal experiences. It is based on the constructivist theory of learning, according to which *understanding* is emphasised rather than algorithms (Carpenter et al., 1994; Davis, Maher, & Noddings, 1990, in Madden et al., 1999, p. 1). Constructivist learning begins with problem solving and "authentic" complex tasks, rather than building up from arithmetic.

1. 15. Summary and relevance of the literature review for the current study

'Pupils' outcomes' that is to say pupils' attainment and progress rates are major indicators associated with the quality of schools, classrooms and the educational system in general. Pupils' outcomes enable researchers to identify structural inequalities embedded in a given school-system. School Effectiveness Research can investigate whether significant differences between schools and/or between classrooms exist, according to a variety of criteria.

Multilevel models are employed in SER to disentangle the impact of schools, classrooms and individual, class-related and school-related characteristics on pupils' outcomes during a given time-period.

Among the studies investigating whether school or classroom effects existed are the American Louisiana School Effectiveness Study (Stringfield (1993), several studies reviewed by Scheerens et al. (1989) and school effectiveness studies falling in the equality of opportunity research tradition, in which also the current study is situated.

One major focus of studies in this tradition is to examine the impact of ethnicity, social class and gender and their interactions on pupils' attainment and progress rates, as these factors are associated with conditions of educational and social disadvantage.

Major representatives of the equality of opportunity research tradition in primary school effectiveness research are Mortimore et al. (1988), Mortimore et al. (1994), Sammons (1995, 1997) Sammons et al. (1993), Thomas (1995), Tizard et al. (1988), Plewis

(1991), Schagen (1994), Brandsma (1993), Hill and Rowe (1996), Strand (1997, 1998, 1999), Brandsma and Knuver (1989).

These studies estimated the impact of ethnic group, sex and social class membership on pupils' attainment and / or progress rates during preprimary or primary school. Many studies co-examined the impact of other variables such as incomplete fluency in English, number of terms in infant school, family size, term of birth, along with contextual school and classroom characteristics such as percentage of pupils with free school meals in the school.

Significant interactions between ethnic group, sex and social class are sometimes reported by these studies (e.g. Tizard et al. (1988), Plewis (1991) and Strand (1999)).

These results are in line with the results of Gilborn and Gipps (1996), or Gilborn and Mirza (2000) and Meighen (1997), who although not adhering to the SER paradigm found that the effects of social class are interlinked with those of ethnicity and gender and therefore should be co-examined.

Some studies (e.g. Sammons, 1995) compare the results obtained across subjects or according to a variety of criteria to identify whether schools / classrooms are consistently effective in two or more subjects or according to more than one criteria. These authors identify complexity in the characterization of effectiveness of classrooms and schools. They ask: effective according to which criteria, effective according to which outcomes, and finally '*effective for whom?*' as Slee et al. (1998) point out.

The latter question is also explored when the differential effectiveness hypothesis is examined. According to this hypothesis, schools or classrooms may produce different value-added results for different categories of pupils (high and low achievers, different ethnic groups or groups defined by minority/majority status or stage of fluency in the native language, or by economic disadvantage or gender). Random coefficients allow the researcher to investigate the differential effectiveness hypothesis.

The Swann (1985) Report showed that ethnic minority pupils in Britain underachieve and argued for equal educational opportunities for minority pupils. Kim-Dong-il (2001) defines underachievement as performance not measuring up to this pupil level of aptitude. The major finding of the Swan Report stressed that multiple explanations are at interplay, and single factor explanations should be avoided.

Extended time for learning characterises most effective international support schemes including those described before (Title I, Success for All, Reading Recovery). Hence, it is intriguing to additionally investigate whether the time offered to support teaching to minority pupils in this Greek setting is adequate for them to cover their learning gaps, which implies to also look into the degree of implementation of Greek support schemes.

2. METHODOLOGY

2. 1. *Aims of the study*

The study initially examines whether school effects or classroom effects exist, in other words whether some of the schools and/or some of the classrooms studied are more effective than others in promoting pupils' learning in language and in mathematics. Also according to the typologies of School Effectiveness studies adopted by Scheerens and Bosker (1997), the current study broadly falls into the '*equality of opportunities*' research tradition; it explores differences in attainment and in progress between pupil groups, defined by ethnicity, social class and gender. More specifically, the study examines whether there is underachievement among foreign/repatriated pupils and pupils from lower social class in terms of their educational outcomes, namely in terms of attainment and progress that pupils make during the last year (year 6) of primary school. The absence of differences between foreign/repatriated pupils' attainment and the attainment of Greek pupils can be perceived as an indicator of successful integration for the former in the Greek school. Or given that several years are needed for foreign/repatriated pupils to become proficient in the Greek language, the absence of difference in attainment between foreign/repatriated and Greek pupils controlling for the years they have spent in Greek primary school can be perceived as another indicator, showing the degree to which these groups have been integrated into the Greek school. In cases where significant differences in terms of absolute attainment score exist, then the existence of positive progress rates for foreign/repatriated groups in relation to the

majority group show that these foreign/repatriated groups are progressing towards the goal of integration.

The absence of differences in attainment between pupils from lower social class and pupils from higher social class can be perceived as an indicator of equal educational opportunities for these pupil groups. If pupils from lower social class are shown to have significantly lower attainment in relation to pupils from higher social class (that they underachieve), the existence of positive progress rates for pupils from lower social class in relation to pupils from higher social class would signify that these pupils are well on the way towards catching up. On the other hand, the existence of negative progress rates in a progress model (the fact that minority pupils progress less than majority pupils) signifies that the gap in educational attainment between pupils from lower and pupils from higher social class widened during the year 1999-2000 when the current study was conducted. If this trend persists, (if pupils from lower social class show negative progress rates during successive years) then their attainment is expected to fall further and further behind in relation to the attainment of pupils from higher social class.

2. 2. Research Design

A longitudinal research design was adopted; pupils were tested twice in basic skills (language and mathematics) at the beginning and at the end of the 6th grade. Individual pupil characteristics were measured on the 1st testing occasion and missing data were filled up on the 2nd testing occasion. Contextual classroom characteristics were calculated on the basis of data collected on the 1st testing occasion. The two measurement occasions were the beginning and the end of the school year 1999-2000.

Data concerning the implementation of the support schemes was collected from principals' questionnaires, which were administered at the end of the school year. '*In the two-phase design the two paradigms are clearly separate*' (Creswell, *op. cit.*, p. 177).

The school effectiveness methodology was adopted to disentangle the contribution of individual pupils' characteristics from the classroom and the school pupils' attend on pupils' attainment and progress and to estimate the effect of contextual classroom characteristics on pupils' progress.

A core issue examined in the current analysis, which is usually dealt with in School Effectiveness studies, is to what extent the school or the classroom that the pupils attend makes a difference to pupils' attainment and/or progress, or whether school and / or classroom effects exist. Furthermore, this study estimates the magnitude of these effects, how much of the total variation in pupils' total scores can be attributed to the fact that pupils are enrolled in schools or in classrooms. The study further looks into the de-

gree to which there is consistency in the way schools/classrooms performed in the two subjects investigated, in language and mathematics, according to a variety of criteria, and more specifically, the following: a) In terms of unadjusted attainment, b) in terms of attainment adjusted for family background, c) in terms of unadjusted progress made within the year the study was conducted, d) in terms of progress adjusted for individual pupil characteristics, e) and finally in terms of progress adjusted for compositional or contextual effects of the classrooms.

School effectiveness methodology can also disentangle the impact of factors operating within classrooms or within schools, as conditions contributing to school or to classroom effectiveness. Dimensions within the school climate and the classroom climate can be considered as examples of such factors, along with pedagogical components or dimensions adopted by teachers in the classrooms; other factors may be the existence or not of a particular intervention scheme in the school, the teachers' attitudes etc. However, this option was not pursued, due to space and time limitations.

Questionnaires that were administered to school principals collected information about the period (in months) when the support schemes were operating in each school, and questionnaires administered to support class teachers collected information about whether the operation of these support schemes was carried out without major interruptions during the school-year or not. Further analysis of the qualitative questionnaires administered to the principals and support class teachers appears in Nikolaou and Kori-laki (2005).

To sum up, the following research questions are examined in the current study:

2. 3. Research questions

1. To what extent does the school or the classroom attended explain differences in attainment and in progress between pupils during the 6th grade? In other words are there any differences in effectiveness between schools or classrooms? What percentage of the total variation in pupils' educational outcomes¹³, can be attributed to the school they are enrolled in and what percentage can be attributed to the classroom?
2. Is there any evidence of differential effectiveness among schools or classrooms? Do schools or classrooms contribute differently to the progress of pupil groups enrolled in them? For example do schools/classrooms contribute more to the progress of initially high achieving pupils versus low achieving pupils? Or do schools/classrooms contribute differently to the progress of pupil groups defined along ethnic, social class and gender lines?
3. To what extent do school and /or classroom residuals derived from value-added (progress or adjusted progress) models differ significantly from those derived from analyses of adjusted attainment models?
4. Are there any attainment differences and/or any differences in progress rates between groups of pupils defined by ethnicity, social class and gender in
a) mathematics and b) language? More specifically are there any differences between foreign/repatriated and native Greek pupil performance; between pupils from

13 In terms of unadjusted attainment, in terms of attainment adjusted for conditions of disadvantage, in terms of unadjusted progress, and finally in terms of progress adjusted for conditions of disadvantage.

higher and from lower social class; and between boys and girls? In other words is there any evidence of underachievement among groups defined by ethnicity, social class and gender at the end of the 6th grade of primary school and in terms of progress made during the 6th grade?

5. Are there any attainment differences and/or any differences in progress rates between groups of pupils defined by nursery school attendance, weekly hours of homework and number of friends in class? Is there any evidence that having friends, doing homework and not having many absences alleviates the impact of ethnicity and social class on pupils' final attainment scores and on their progress rates?
6. Are there any significant differences in attainment and in progress rates between foreign/repatriated pupils from various ethnic origins? Namely are there any differences between Albanian pupils, pupils from the democracies of the former USSR, and pupils from 'other' ethnic groups (which are grouped in one category) and the majority group with respect to their performance in language and in mathematics?
7. Are there any significant interaction terms between all individual pupil characteristics examined as well as between them and initial attainment score, which have an additional impact on pupils' attainment and progress rates?
8. Is there any evidence that contextual or compositional effects of the school or the classroom pupils attend have an impact on pupils' progress rates over and above the effects associated with individual pupils' prior attainment or background during the final year of primary school?

9. According to principals' accounts to what degree has support teaching to foreign/repatriated pupils been implemented so far? Have support schemes currently available in the Greek setting been implemented in a consistent way, or have they been poorly implemented?

2. 4. Analysis techniques employed in the quantitative part

In order to address the above research questions, the researcher applied ‘hierarchical linear modelling’ broadly known as ‘multilevel modelling’ to the sample.

Teddlie et al. (2002) refer to the decision to gather data at the multiple levels of the classroom and the school ‘*which has become axiomatic in SER over the past decade as both theory (e.g. Creemers, 1994; Scheerens and Creemers, 1989) and mathematical models capable of analysing multiple levels of data have been developed (e.g. Atkin and Longford, 1986; Bryk and Raudenbush, 1992; Goldstein, 1995)*’.

Multilevel modelling aims to disentangle the effects of individual pupils’ characteristics on pupils’ attainment and progress rates from the effects of belonging to a particular classroom and/or school and this allows school effects to be explicitly modelled. The net impact of the school or the classroom in terms of progress constitutes the ‘value added’ of the school or the classroom. Goldstein (1997, p. 376) places among the minimum conditions for satisfactory inference about individual school performances that a school effectiveness study should be longitudinal so that pre-existing differences and subsequent contingent events among institutions are taken into account; also a proper multilevel analysis should be undertaken so that statistical inferences are valid and in particular the question of whether a school is differentially effective is explored. Goldstein (1995, 2003) provides an in-depth statistical and theoretical account of multi-level modelling techniques.

According to Teddlie et al. (2000), in order to study change over time, it is necessary to study schools longitudinally. By adopting a longitudinal design the estimation of pro-

gress rates for various groups of pupils defined by ethnicity, social class, gender etc. can be estimated in relation to the control group. This comparison is particularly relevant in the case of second language learners in relation to the majority group. In addition, classroom or school comparisons in terms of progress are enabled.

As Thomas and Collier (1997, p. 72) point out:

‘Some school-based questions that focus on the present status of selected variables (e.g. attendance, disciplinary actions, and current achievement levels) require cross-sectional data and can be addressed using a short-term outlook. However, the impact of appropriate education for (second) language learners requires a long-term look at trend data, and a continuous monitoring of the progress that students make over a number of years. For these questions, only longitudinal data will do’.

A longitudinal design in its simplest form assesses pupils’ attainment levels at two time-points: one at the beginning and one at the end of the assessment period. In more elaborated forms of SER multiple assessment points are employed. On top of initial attainment score, individual pupils’ characteristics, as well as characteristics of the classroom and/or the school pupils enrolled, are incorporated. Examples of individual pupils’ characteristics are social class, ethnicity and gender. Variables that can be measured at the school level such as school’s organisational climate or the provision of support teaching to disadvantaged pupil groups, administrative leadership, curriculum content, utilization of resources and variables that can be measured at the classroom level, which represent classroom processes and instruction such as time on task, time for management, preferred teaching styles and classroom climate, have not been included

in the current study's design. Only contextual classroom effects have been introduced in the current analysis as variables at the classroom level.

Multilevel Modelling allows for comparisons between schools and / or classrooms to be carried out both simultaneously (using a cross-sectional design) and over a period of time, (using a longitudinal design). When a cross-sectional design is used, the impact of individual pupil, as well as classroom level and school level variables on pupils' final attainment score is estimated, adjusting for other explanatory variables. When longitudinal design is used, the impact of these variables on pupils' progress rates is estimated. The longitudinal design enables the final attainment score for each individual pupil to be adjusted for his/her initial attainment score.

Multilevel modelling also enables comparisons between classrooms and/or schools in terms of unadjusted attainment and/or unadjusted progress or in terms of adjusted attainment and/or adjusted progress (through the examination of the rankings of classroom residuals derived from the respective model). When no explanatory variables are introduced into a null model (that includes only pupils' final attainment score as a dependent variable), unadjusted comparisons between classrooms in terms of average final attainment scores are undertaken. When variables representing individual pupils' characteristics are introduced into an absolute attainment model, adjusted comparisons in terms of final attainment score adjusted for these characteristics between classrooms are undertaken. In contrast, when classroom residuals are derived through a progress model that adjusts for initial attainment score, comparisons between classrooms in terms of progress are undertaken. When classroom residuals are derived through an adjusted progress model that adjusts for pupils' initial attainment score and for variables

representing individual pupils' characteristics, adjusted comparisons in terms of progress between classrooms serving pupils with similar characteristics are undertaken. Finally, when a progress model adjusts for pupils' initial attainment score along with individual pupils' characteristics and contextual classroom characteristics, comparisons in terms of adjusted progress between classrooms similar both in terms of pupils' composition in the class and also in terms of individual pupils' background characteristics are undertaken.

Unadjusted classroom comparisons are called classroom 'type A effects' according to Raudenbush and Willms's (1995, p. 309) distinction, while classroom comparisons adjusted for pupils' prior attainment scores are called 'type B effects'. Type A effects reflect the classroom's attainment irrespective of the circumstances that led to this attainment, which may or may not be within the realm of classroom or school. Teachers' efforts, principals' initiatives, a positive classroom climate, along with contextual factors in the classroom and community's efforts, may all contribute to the creation of a classroom and school environment which is more or less conducive to learning. Type B effects, on the other hand, are designed to isolate the effect of school or classroom practice from other positive or negative contextual effects, that jointly contribute to school or classroom final performance level. Raudenbush and Willms (1995b) distinguish this practice from school context, which includes school-level factors that are exogenous to the practices of the school's administrators and teachers. Such contextual factors include the social and economic characteristics of the community in which the school is located and the demographic composition of the student body.

Type B effects correspond to ‘progress’ or to ‘adjusted progress’ models. Type B effects can be employed for purposes of school/classroom accountability, whereas type A effects can be employed for intervention purposes on the basis of equal educational opportunities.

While differences in absolute attainment between pupil groups are estimated using a cross-sectional design, differences in pupils’ progress (during the school-year 1999-2000) are estimated using a longitudinal design. Both cross-sectional and progress models are needed in order to monitor the performance of foreign/repatriated pupils, or of pupils from lower social class background. While absolute attainment models indicate the relative attainment of given pupil groups at a certain time-point, progress models indicate differences in progress between groups; progress differences can serve as a warning of whether attainment gaps identified are going to be bridged in the future or not.

Furthermore, combinations of individual pupils’ characteristics are introduced in absolute attainment models (that do not adjust for pupils’ initial attainment score) or in progress models (that adjust for pupils’ initial attainment score). Absolute attainment models correspond to a cross-sectional research design. Progress models correspond to a longitudinal design.

When combinations of individual pupils’ characteristics are introduced in an absolute attainment model or in a progress model, the impact of each individual characteristic on pupils’ attainment or on pupils’ progress is estimated controlling for all the remaining characteristics in the model. The inclusion of additional variables in a model has the

effect of reducing the coefficients of other variables measuring individual characteristics that have been introduced previously. The reduction of these coefficients can be perceived as the reduction of the impact of these variables in the presence of a recently introduced variable. In other words the introduction of an additional variable often alleviates the impact of variables introduced previously in a given model. Haque and Bell's (2001) study revealed that differences in attainment and in progress that were identified between pupils from minority ethnic background are substantially reduced when background factors (other than ethnic origin) are taken into account.

In order to assess the net impact of a given variable on pupils' attainment and progress, nested absolute attainment or nested progress models are constructed. These models start by examining the impact that a single individual characteristic on pupils' final attainment scores or on pupils' progress rates. Subsequent models control for additional characteristics, having controlled for the initial characteristic of interest. For example, if one aims to build models that estimate the impact of gender on pupils' final attainment score, then the researcher starts model building by creating an absolute attainment model linking pupils' final attainment-score with gender. Then subsequent models adjust for gender and another variable. Two variable combinations are produced (e.g. gender and social class, gender and ethnicity, gender and homework etc.). Subsequently the analyst examines how the coefficient of gender increases or decreases in the presence of other variables in comparison with the initial absolute attainment model that employs gender as a single explanatory variable.

The existence of any significant interaction effects among all variables measuring individual pupils' characteristics in terms of pupils' attainment and in terms of pupils' pro-

gress is also investigated. Interaction effects are included in both absolute attainment and in progress models already including the corresponding main effects terms.

2. 5. Describing pupils' characteristics

The rationale of the inclusion of pupils' background characteristics is discussed below.

Ethnicity: One question addressed to all pupils in any given class before they began to complete the performance-based tests, was whether they had come from another country. Teachers then confirmed pupils' answers and in some cases pinpointed pupils who were afraid to reveal their foreign identity. Hence the ethnicity variable is considered to be very reliable. To start with, foreign and repatriated pupils were grouped together in the same category in the analysis. In such a case their joint performance as a foreign/repatriated group is compared to the performance of Greek pupils. Foreign pupils are those with foreign citizenship, whereas repatriated pupils are pupils who initially were of Greek origin, but may have been in the foreign country (usually in the democracies of the former USSR) all their lives, while their families have been there for many generations. While in some cases foreign and repatriated pupils have a very limited command of the Greek language, often they do not know any Greek at all. Subsequently the educational outcomes for three separate ethnic groups were estimated in relation to the outcomes of Greek pupils. Namely the outcomes for '*Albanian*' pupils, for pupils from the democracies of '*the former USSR*', and for pupils from '*other*' ethnic origins are estimated. Albanian pupils made up 8 % of the current sample (152 pupils), pupils from the democracies of the former USSR 3 % (56 pupils) and pupils grouped in the '*other*' category 1.7% of the current sample (33 pupils). In the '*other*' ethnic group pupils belonging to all the remaining ethnic groups are grouped together. The category '*pupils from the former USSR*' includes pupils from *Russia, Kazakhstan,*

Uzbekistan, Armenia, The Ukraine, Moldavia. The category ‘*other*’ includes pupils from *the U.S.A., Canada, Mexico, the Dominican Republic, Brazil, Syria, Iraq, Sri-Lanka, The Philippines, Australia, Japan* and finally *Greek-Muslim pupils*. Greek Muslim pupils have Greek citizenship, but speak Turkish as a first language. The above ethnic groups had to be grouped together in the ‘*Other*’ category, as the sample size of some of the minority ethnic groups (e.g. Philippino or Polish) proved to be too small for detailed multilevel analyses. ‘*It is common practice in statistical studies to combine small and apparently similar groups to form a category of reasonable size in order to reduce the vagaries of small sample sizes*’ (Strand, 1999, p. 197). More specifically, the following ethnic groups were encountered:

Table 2. 1: Pupils' citizenship

	Frequency	Percent (%)
<i>Greek</i>	<i>1558</i>	<i>87.2</i>
<i>Foreign</i>	<i>195</i>	<i>11</i>
German	3	.2
British	1	.1
Polish	2	.1
Roumanian	2	.1
Bulgarian	2	.1
Albanian	129	7.2
Ukrainian	2	.1
Russian	28	1.6
Moldavian	2	.1
Armenian	4	.2
Kazakstani	4	.2
Ouzbekistani	1	.1
Canadian	1	.1
Mexican	1	.1
St Dominican	1	.1
Brazilian	1	.1
Syrian	1	.1
Iraqi	2	.1
Philippino	1	.1
Japanese	1	.1
Australian	1	.1
Muslim	5	.3
Total	1753	98.2
Missing	33	1.8
	1786	100.0

Years in Greek school: This variable shows the number of school years foreign / repatriated pupils had already spent in the Greek school when the study was conducted. It was collected through pupils' themselves but it was confirmed through their teachers' accounts and therefore it is considered to be reliable. The more school years spent in the Greek school the better foreign/repatriated pupils' ability is to communicate, understand, read and write the Greek language, and also the more familiar they are expected to become in relation to the school's culture, organisation and environment.

Thomas & Collier (1997, p. 37) found that the number of years of exposure to the English language is a strong predictor of second language learners' long-term academic achievement. All groups, whatever their circumstances, demonstrate growth in development of their second language (Greek) for each additional year of exposure to this second language.

Haque & Bell (2001) conducted a school effectiveness study in secondary schools, which identified that recency of arrival in the host country had a large negative effect on students' attainment. Moreover, when in the same model, recent arrival reduced the size of ethnic origin parameters (differences) substantially. In the Netherlands, Driessen (1993, 1995 in Driessen and Mulder, 1999) showed that minority pupils perform better the longer they have been in the host country.

Pre-school education: Pupils reported whether they had been enrolled in nursery school or not. As the data was collected from pupils' self-reported accounts, some caveats about the reliability of this variable should be included. The variable takes the value of 1 if a pupil has been enrolled in nursery school and 0 if the pupil has not been

enrolled in nursery school. It has been found that children who have attended nursery school are at an advantage when starting real school (Hutchison, 1993). In the attainment of foreign/and repatriated pupils and of pupils from disadvantaged backgrounds in general nursery school education has been shown to play an important role both by the English Plowden report (Central Advisory Council for Education, 1967), and by the Headstart programme in the USA (U.S. Department of Health and Human Services, 2003). According to the latter study children who recognize their letters, who are read to at least three times a week, who recognize basic numbers and shapes, and who demonstrate an understanding of the mathematical concept of relative size in kindergarten have a difference of one year of reading age in reading skills at the end of the first grade from children who do not have this background. The report claims that this is true regardless of family income, race or ethnicity.

Gender: Pupils' gender is coded with 1 if the pupil is a boy and with 0 if the pupil is a girl. Pupils' gender is the most reliable variable, as pupils' name reveals whether the pupil is a boy or a girl. The impact of gender on pupils' attainment and progress rates has been shown in Mortimore et al. (1988), Strand (1997, 1999), and Thomas (1995). Gender has also been found to interact with other variables such as ethnicity and jointly to shape pupils' progress rates (Strand, 1999).

Friends: This variable measures the number of pupils who named each pupil as their friend in the same classroom. In the testing occasion the researcher requested each pupil to name three or more pupils in the same classroom, whom s/he considered to be his/her friends. According to pupils' responses, classroom *sociograms* were drawn connecting each pupil's name with arrows indicating the names of pupils who named him/her as

their friend. Finally, the arrows pointing to each pupil in the class were added up and formed a number next to each pupil's name. The coefficient of having many *friends* can be perceived as an index of the degree of integration of a given pupil in his/her classroom. '*Sociograms are relatively non-reactive instruments designed to measure the social structure of a group and to assess the social status of each individual in the group*' (e.g. Borg and Gall; 1989; Moreno, 1953, in Reynolds and Teddlie, 2000). As friendships pupils make are constantly evolving caveats about the reliability of this variable should be included as well.

The number of friends a foreign/repatriated pupil has is an important predictor of the pupil's social integration; it is known that social integration and interaction with native peers is a factor helping the pupil to acquire the second language. Thomas & Collier (1997, p. 51) judge that: '*Language learners need access to meaningful interaction with native speaking peers in a supportive environment. Same age peers are a crucial source of L2 (second language) support*'. In addition, '*some children in some cultures are more accustomed to learning from peers than from adults*' (in McLaughlin, 1992, p. 6).

Friendships are formed on a social class basis as well. It is a familiar argument in stratification theory that persons sharing a similar social position, in terms of social class or status group membership, are more likely to interact socially on the basis of equality with members of the same group than with members of other groups (Prandy, 2002).

Peer relations contribute substantially to both social and cognitive development (Hartup, Willard, 1992, p. 1). Having many friends is the best childhood predictor of adult adaptation. The author perceives friends as emotional resources, both for having

fun and adapting to stress, as cognitive resources for problem-solving and knowledge acquisition and as contexts in which basic social skills are acquired or elaborated, and as forerunners for subsequent relationships. Friends are similar to each other in developmental status, engaging each other mostly in play and socializing.

From a psychological perspective, friendships may protect children from the adverse effects of negative events, such as family conflict, terminal illness, parents' unemployment, and school failure, easing the stress involved.

From a cognitive perspective, friends often teach one another effectively. Peer teaching occurs as peer tutoring, as collaborative learning, as peer modelling (Hartup, Willard, *op. cit.*). Pupils reported to receive low social support or none at all were reported to have lower attendance, to spend less time studying, to have fewer friends and pro-social behaviours and to have less ability to overcome school problems (Rosenfeld et al., 1998).

The number of friends a pupil has, as well as the self-reported popularity were employed as predictors of peer victimisation among 52 seventh and eighth graders. Pupils with high popularity were victimised less often than pupils perceived as unpopular (Coleman and Byrd, 2003).

Fewer and less optimal peer relationships and friends are also associated with depressed adolescents, who are less popular (Field et al., 2001). Isolated youth were also reported to demonstrate more shyness, greater feelings of alienation, and lower social acceptance than did integrated youths (Tani et al., 2001).

The number of friends a boy has early on is associated with adolescent delinquency. Having poor relations before the age of 13 is associated with early adolescent onset in offending (Stattin et al., 1995). These boys often dropped out from school. By contrast, boys with middle and late adolescent onset in offending displayed less evidence of poor peer relations in early grades and had many close friends at school and same age friends. Pettit et al. (1999) established that unsupervised peer contact in the after-school hours was examined as a risk factor in the development of externalising problems in a longitudinal sample of adolescents aged 12-13.

Absences: The variable '*absences*' represents the number of days each pupil had been absent from the school, during the school year. The more days a pupil has been absent, the less time the pupil has been exposed to instruction. Rutter et al. (1979) also investigated attendance rates as a dependent variable in their study. The investigators report large differences in attendance rates across schools, even after variations in school input were controlled for.

Caveats about the reliability of absences should be included, as pupils themselves reported the days they had been absent during the school year and hence some of them might not remember accurately, and as the time that the final measurement was taken in various schools ranged from mid-April to the beginning of June. Pupils enrolled in schools measured early might have underestimated the actual number of days they had been absent during the year.

Homework: Homework is described as '*work set by teachers that pupils are expected to complete in out of school hours*' (Cooper, 1989a and 1989b in Sharp et al., 2001, p.

7). Pupils were asked to indicate how many hours of homework they used to study during a typical week, and this number of weekly hours of homework is registered. This amount of homework undertaken is related to the 'effective learning time' according to Scheerens's (1992) review of the research evidence on effective teacher characteristics.

Scheerens and Bosker (1997) in their review of effective school correlates in SER estimated that the mean effect size of homework across all studies was $z=0.0574$, significant at the 5% level. Cooper (1989) has reviewed homework studies, and cited 14 studies (out of a total of 20 studies) where interventions showed a positive relationship between homework and attainment. *'Among the suggested positive effects of homework, the most obvious is that it will have an immediate impact on the retention and understanding of the material it covers'* (Comer, 1989, p. 86). However, Comer continues, *'children from poorer homes will have more difficulty completing assignments than their middle class counterparts ...Homework it was agreed is not the great equaliser'* (Comer, 1989, p. 87).

Holmes et al (1989) investigated the relationship between time pupils spend on homework according to pupils' accounts and their level of achievement in public examinations and concluded that *'levels of time on homework had a fairly strong association with academic achievement'* (Holmes et al., 1989, p. 36). This relationship was maintained even after controlling for other variables such as ability and family background.

Cawelti (1995) synthesized more than a dozen studies on the effects of *homework* in various subjects. These studies showed that *'the assignment and the completion of homework yield positive effects on academic achievement. The effects are almost tri-*

pled when the teachers take time to grade the works, make corrections and specific comments on improvements that can be made, and discuss problems and remedies with individual students or the whole class' (Cawelti, *op. cit.*, p. 20).

Paschal et al. (1984) in a meta-analysis, also concluded that homework had a positive effect on academic achievement, especially when commented upon or graded.

Van der Werf (1995) found that the amount of homework showed the relatively largest effect on pupils' achievement in a school effectiveness primary school study conducted in 1988 in the Netherlands (cited in Scheerens & Creemers, 1999).

According to Rutter (1980) homework was found to be an important within – school factor determining high levels of effectiveness. Rutter reported that successful schools in the U.K. tended to make good use of homework, to set clear academic goals and to have an atmosphere of confidence in the work of their pupils' capacities (reported in Reynolds, 1992, p. 9).

Farrow et al. (1999) found support for the view that the '*amount of homework undertaken at year 6 was positively associated with pupils' performance in mathematics and science*'.

In OFSTED (1995, p. 2) homework appears to have '*the potential to raise standards, extend coverage of the curriculum, allow more effective use of lesson times and improve pupils' study skills and attitudes to learning*'. This report suggests that there was '*considerable variation among schools in the amount of homework set for each year group*' in England.

As pupils' own accounts about the amount of homework undertaken during a typical week were recorded, caveats about the reliability of the homework variable should be included as well. Some pupils might tend to inflate whereas others may underestimate the amount of homework they undertake during a typical week.

Social class: Pupils were asked to designate the current occupation of their father and of their mother, before starting to complete the performance-based tests. The researcher and the research assistants clarified the pupil's answer if the response was too vague. For example, if the pupil answered 'my father works in a shop', then the researchers prompted, 'Is he the owner in the shop or is he an employee in the shop?', so that a more accurate categorisation of the pupils' answers to Goldthorpe social class categories was made possible. However, in spite of the researcher's and the research assistants' efforts, mistakes in the accurate classification of parents' occupation are difficult to avoid given that the data comes from pupils' own accounts. Data of better quality would have been produced if parents' own descriptions of their occupations had been collected at the time. Hence this data may be of somewhat reduced reliability. Coe and Fitz-Gibbon (1998, cited in Hutchison, 2003) in a far-reaching critique of the school effectiveness research paradigm have commented that unreliability in control variables may lead to bias in the school effectiveness measures. In the UK, School Effectiveness studies have repeatedly employed 'free school meals' as a proxy variable for social class (for example, Strand, 1997, 1999) and found that pupils with free school meal entitlement have lower attainment and progress rates than pupils without free school meal entitlement. The free school meals variable has also been found to interact with other

variables such as ethnicity and jointly to shape pupils' attainment and progress rates (Strand, 1999).

In the current study, in order to operationalise pupils' social class, the Goldthorpe social class schema was adopted. This finely differentiated schema was chosen in order to overcome one acknowledged weakness encountered in many SER studies, namely the fact that *'social background measures are not usually finely differentiated'* (Gray and Wilcox, 1995, p. 126). This schema originally employed 11 categories, even though only 9 of them were encountered among the pupils included in the current sample.

'The aim of the class schema is to differentiate positions within labour markets and production units and more specifically ...to differentiate such positions in terms of the employment relations that they entail' (Erikson and Goldthorpe 1992b: 37, in Breen and Rottman, 1995, p. 70).

Breen and Rottman describe Goldthorpe social class classification in the following manner:

'In focusing employment relations the first distinction is made between employees, the self-employed and employees. Among the self-employed a sectoral distinction is made between farmers (class IVc in the schema) and non-agricultural self-employment (class IVb). Within the group of employees classes are further defined on the basis of the employment relationship they enjoy. Here a chief distinction is between occupations that are regulated by a service relationship and those based on a labour contract. So the service class (classes I and II in the schema) enjoy, as their name suggests, a service relationship, while the manual classes (VI and VII) are those with a labour contract

relationship with their employer. Between these two extremes lie the aptly named intermediate classes (III and V)'.

Erikson (1984) disentangles the issue 'whose parent's occupation should determine the child's social class' by stating:

'If the mother does not work (is a housewife), she and her children will be assigned to the social class of her husband. In the case of dual-earner families, such as those in which both the father and the mother of the child are working, the 'dominance' approach should be used'.

It is suggested that each child be assigned to a class based on occupation of whichever person is considered dominant. Erikson takes the dominant person to be whichever parent has the occupation with the greater impact on the family's life chances. So, non-manual dominates manual work; professional employment dominates self-employment; the latter in turn dominates work as an employee, and so on.

Table 2. 2: Detailed Goldthorpe social class schema

Goldthorpe class designation	Description	Employment relations
I	High grade professionals, administrators and officials; managers in large industrial establishments; large proprietors	Employer or service relationship
II	Lower grade professionals, administrators and officials; higher grade technicians; managers in small industrial establishments; supervisors of non-manual employees	Service relationship
IIIa	Routine non manual employees, higher grade (administration and commerce)	Intermediate
IIIb	Routine non manual employees, lower grade (sales and services)	Intermediate
IVa	Small proprietors, artisans, etc., with employees	Employer
IVb	Small proprietors, artisans, etc., without employees	Self-employed
IVc	Farmers and smallholders; other self-employed workers in primary production	Employer or self-employed (Not encountered)
V	Lower grade technicians; supervisors of manual workers	Intermediate
VI	Skilled manual workers	Labour contract
VIIa	Semi- and unskilled manual workers (not in agriculture, etc.)	Labour contract
VIIb	Agricultural and other workers in primary production	Labour contract

Using the Goldthorpe social class classification schema, one can estimate the impact of belonging to a specific social class category in relation to a base category. For the current study the base category is taken to be the highest category, namely category I, including ‘*High grade professionals, administrators and officials; managers in large industrial establishments; large proprietors*’.

From the above 11 categories, categories IVb and VIIb, namely the categories ‘*farmers and smallholders; other self-employed workers in primary production*’ and ‘*Agricultural and other workers in primary production*’ were not found among pupils’ parental

occupations. As Piraeus is an urban area, it is very rare to find agricultural jobs among the parental occupations. In the case of this data set, there were none. Consequently, nine only of the original 11 categories in the Goldthorpe scheme were employed. Subsequently, a dummy (dichotomous) variable was created based on the significant and non-significant coefficients that came out of the absolute attainment models employing the detailed nine-category Goldthorpe social class scheme.

Mother's educational level: Pupils were asked to provide the highest educational level completed by their father and by their mother, before starting to complete the performance-based tests. If the children did not understand that, they were prompted to answer the following question: *'Did your mother complete primary education, secondary education-lower level (gymnasio), secondary education-upper level (lykeio)? Do your parents have any secondary technical qualification, University degree or degree from some tertiary technical school?'*, followed up by the question *'If so, which one?'*. Yet, as children might not have known the exact grade after which each of their parents left school, caveats about the reliability of this data should be made as well.

Increased educational attainment of parents is associated with pupils' positive outcomes in school readiness (Nord et al., 1999). These authors characterised less than high school education as a risk factor in the development of pupils' literacy skills, as the higher the educational level a mother has successfully completed, the more likely she is to read stories to her child, to tell a story, or to visit the library with the child. Children whose mothers had less than a high school education were less likely to engage in literacy activities than the ones explained previously.

In their overview of the determinants of children's attainments, Havenman and Wolfe (1995) conclude that the most fundamental factor describing children's educational attainment is the human capital of their parents, typically measured by the number of years of schooling attained. The human capital of the mother is usually more closely related to the attainment of the child than is that of the father. Children of better-educated immigrants have higher education and earn higher wages. Couch and Dunn (1997) found that German children's education has very weak correlations with their mothers, whereas in the US the correlations are of the same magnitude as the correlations with father's education. Dearden, Machin and Reed (1997) found that the education of both parents has a strong impact on the education of their children but, whereas father's education is more important for sons, mother's education is more important for daughters. Van Ours and Veenman (2001, p. 15) investigated whether the level of education of the father is more or less important than the educational level of the mother, but they could find no difference in the relative importance.

Also, in a school effectiveness study carried out recently in Indonesia, Kaluge (1998) found that *father's education* was a significant predictor of educational attainment of the child. The higher the *education* the *father* had, the better the language attainment of the child (Kaluge *op. cit.*, p. 127).

The Swedish National Agency for Education (2001) carried out a study that intended to contextualise school achievement scores in relation to pupils' background factors, in order to establish a basis of comparisons between schools and between municipalities. In analysing the impact of background factors on achievement value, the analysis

showed a positive connection with parental educational level and achievement – the higher the educational level, the better the average achievement.

Thomas & Collier (1997, p. 39) reported that '*The amount of formal schooling parents have completed can be a very significant predictor of their children*'.

Kerckhoff et al. (1973) investigated the accuracy of pupils' reports on parental status and occupation in relation to those made by the parents themselves. Sixth, ninth and twelfth graders were included in the sample. Both parents were asked how far they had gone in school and the father was asked about his current occupation. Then their children were asked to answer the same questions. The authors reported a serious loss due to pupils' non-response. They also found that the degree of agreement between parent and son increased the higher the son's grade. This finding was true for both white and black children. Hence, they concluded that it is possible to obtain high response rates from young children without also reducing the accuracy of the data. However, a degree of wariness is needed because the accuracy of sons' ratings is a function of age.

The authors identified only a slight tendency to upgrade parents' characteristics, primarily in sixth graders' reports of mother's education. This finding can be contrasted with that of St John's (1970) according to whom sixth grade boys are more likely to upgrade father's educational level than that of mother's. However, both these authors identified no substantial tendency for children to upgrade family social status. In addition, no racial difference in upgrading was identified.

In the current study mother's education variable has been initially employed as a dummy one with six categories in an absolute attainment model, and it was ascertained

that its coefficients were in an ascending order. Hence, this variable has also been employed as ordinal, taking the values 0 to 6. The variable takes the value of zero if the pupil's mother is completely illiterate, up to six if the mother has successfully graduated from a university.

Next, the categories 0-3 were collapsed to form the base category in a new dichotomous variable. The latter variable measures the impact of the mother having completed compulsory education, which corresponds to nine years of schooling, or higher than compulsory education, over the mother not having completed compulsory education.

Given that pupils' self-reported accounts were used to derive information about their parents' occupation and their mothers' highest educational level reached, some caution should be exercised in relation to the reliability and the validity of these two variables, in addition to reasons previously discussed. This option was undertaken as relevant information was not readily available in the schools' registration records.

2. 6. Structure and properties of the performance-based tests employed in the current study¹⁴ and threats to validity associated with the test construction

According to Anastasi (1988), who adheres to the classical test theory paradigm, curriculum based tests designed to measure pupils' attainment in a given subject on two different occasions should have similar structure, but different content. The tests employed to measure pupils' attainment in the current study sampled contents included in the authorised curriculum. These curriculum-based tests measured pupils' performance in language and in mathematics and they were administered at the beginning (October – December 1999) and at the end (April-June 2000) of the school year. These tests were piloted and constructed by the researcher¹⁵ during the previous year (1988-1999). These tests were administered to both Greek and foreign/repatriated pupils enrolled in the 6th grade. The tests administered at the beginning of the school year were designed to map out curricular dimensions of language and mathematics taught in the 5th grade, whereas those administered at the end were constructed to map out curricular dimensions taught in the 6th grade, defined in pupils' authorised textbooks. Although tests administered at the beginning and at the end of the school year sample different curricula, they have a similar structure, in order that the weight given to different curricular dimensions be similar across grades. For example, in the case of mathematics similar weight is given to calculus, problem solving and geometry across grades. Experts in the field of cognitive psychology in mathematics were consulted in order to define which objectives were core and which ones were peripheral in pupils' future mathematics learning. These

14 A thorough discussion about the tests' structure and properties, and the reliability issue is included in the 1st appendix referring to test construction.

15 Details about the properties and the rationale of performance-based tests can be found in the appendix.

experts were consulted after the pilot, in order to adjust the difficulty level of the test items. Greek language teachers and Greek language experts were consulted as well. The fact that the tests were constructed so as to map out curricular dimensions in authorised textbooks along with the fact that experts in cognitive psychology hierarchised these curricular goals of the textbooks guarantees content validity.

Since different tests were used at the beginning and at the end of the school year, an important threat to tests' interval validity was circumvented, namely to increase pupils' familiarity with the tests employed and hence artificially increase pupils' outcomes on the second testing occasion. Administration of similar or equivalent tests in structure might *sensitise* the pupils, as pupils on the second testing occasion are better trained to the test format. Tashakkori and Teddlie (1998) refer to this as a pre-testing or carryover effect. These authors add that a difference (or lack of difference) between the pre-test and the post-test might stem from increasing pupil familiarity with the test, rather than from their exposure to other (independent) intervening variables.

The first requirement of a high-quality, or 'good' test is that the test possess what is called 'reliability' (Walsh and Betz, 1985, p. 47). The reliability coefficient of the mathematics tests measuring pupils' performance at the beginning of the school year was estimated to be 0.88 and 0.92 for A and B parts of tests respectively. The reliability coefficient for the mathematics tests measuring pupils' performance at the end of the school year was estimated to be 0.89 for both A and B parts. The reliability coefficient of the language test referring to the beginning of the school year was estimated to be 0.94 whereas the reliability coefficient of the language test referring to the end of the school year was estimated to be 0.93. According to Rudner and Schafer (2001), most

large-scale tests report reliability coefficients that exceed 0.80 and often exceed 0.90. Anastasi (1988, p. 115) states that desirable values for reliability coefficients fall in the 0.80s or 0.90s. Thus, typical reliability coefficients were obtained for all six tests employed. These curriculum-based tests were designed to measure performance. They were curriculum-based tests, named performance-based tests as well, '*designed to measure the effects of an instructional programme*' (Anastasi, 1988, p. 411) and not tests measuring linguistic or mathematical ability. According to Anastasi, (*op. cit.*, p. 412), while achievement tests measure the effects of a relatively standardised set of experiences, such as schooling, aptitude test performance reflects the cumulative influence of a multiplicity of experiences in daily living. A criticism made by Gandara et al. (2000, p. 28) to the employment of current tests in California for accountability purposes is that 'the existing test (SAT 9) seriously confounds academic ability with English competence'.

These tests were constructed and piloted one school year before the main study (in 1988-1999) by the researcher herself. They were administered to Greek, foreign and repatriated pupils enrolled in the 6th grade during the beginning (October –December 1999) and the end of the school year (April-June 2000).

A test with a large amount of measurement error does not measure accurately what it purports to measure and therefore scores derived through this instrument can bias subsequent analyses. Such a test has low reliability and low internal validity. This is referred to as 'instrumentation error' in Tashakkori and Teddlie (*op. cit.*). In cases where tests with high measurement error are employed, differences between pupils' initial and

final attainment scores might result from error and they might underestimate or overestimate pupils' progress.

2. 7. Other factors that may reduce the validity of the current study

Pupils' *mobility* between individual schools may distort the findings of a longitudinal study, according to Tashakkori and Teddlie (*op. cit.*). This threat to internal validity is referred to as subject attrition. In the current study pupils' mobility between individual schools was limited. However, the fact that some of the pupils were absent on one testing occasion might have a distorting effect on the estimation of classroom effects and it might contribute to the underestimation of progress differences for different groups of pupils. It is likely that some low achieving pupils avoided the final testing occasion for fear that this test might negatively influence their final grade.

2. 8. Sample

According to data from the National Statistical Service of Greece (NSSG), during the school year 1999-2000 in Greece there were 5,980 state schools with 43,694 teachers and 597,820 pupils. In the city of Piraeus, in which the school effectiveness study was conducted, there were 351 state schools with 54,718 pupils. Given that the tables published did not further classify by grade, the number of pupils enrolled in the sixth grade was estimated by dividing the total number of pupils by six. If we consider that the Greek primary school is comprised of six grades, the number of pupils enrolled in the

6th grade in all Greek state primary schools should be approximately 99,636 pupils, whereas the number of pupils enrolled in schools belonging to the district of Piraeus should be approximately 9,120 pupils.

The selection of schools that took part in the study was made from a sampling frame provided by the Greek Ministry of Education. Among the schools that had six or more classrooms in the 6th grade 58 schools with 83 6th grade classrooms were randomly selected for the purposes of the current study.

At the level of the individual pupil, the sample is a clustered one, since all sixth grade pupils enrolled in all sections in a selected school were tested, that is to say 1858 pupils who were enrolled in these 83 sixth -grade classrooms in the afore-mentioned 58 schools. Thus the sampling fraction that has been employed according to the initial sampling design is 20.3%. From these schools 53 remained in the study on both testing occasions, whereas five dropped out on the second testing occasion. From the remaining 53 schools, 27 had two or more classrooms, while 26 had only one classroom.

For mathematics 1514 pupils were assessed on both testing occasions. Thus for progress models the sampling fraction was 16.6%, which is considered to be sufficiently large. However, the findings might have been slightly modified if all schools had remained in the sample or if all pupils who were tested on the first testing occasion were also tested on the second testing occasion. Hence caution might be exercised in the interpretation of the findings.

It is very important that the sample of schools, classrooms and pupils taking part in the study be sufficiently large. In most multi-level modelling studies, the sample size of the

level-2 units varies between 20 and 200. '*Large samples are preferable to increase the precision of parameter estimates i.e. to obtain tight confidence intervals around the parameter estimates*' (Snijders and Bosker, 1999, p. 23). These authors suggested that large sample sizes of 30 level two units or more should be employed if precise estimation of parameters included in models with random slopes is required.

Also Maas and Hox (2002) investigated the question of what constitutes a sufficient sample size for accurate estimation in multilevel analysis. Their results showed that for studies employing two levels sample sizes that can lead to biased estimates of the second level standard errors are 50 units or less. In all other simulated conditions the estimates of both the regression coefficients, the variance components and the standard errors are unbiased and accurate. Hox (1998) cited in Sastry et al. (2003, p. 7) suggested that a good rule of thumb for estimating models with random slopes is to have about 50 groups with 20 individuals per group.

Afshartus (1995) study investigated the minimum number of groups needed to obtain unbiased, stable, and efficient parameter estimates, holding group size fixed. He found that as few as 40 groups were sufficient for estimating regression coefficients but that as many as 320 groups were needed to estimate variance components.

Therefore the main effects of variables included in absolute attainment and progress models in the current study can be considered as unbiased since the study employed 83 classrooms with an average class size of 22 pupils in 55 schools. The only caution should refer to the estimation of the random effects or variance components, since ac-

according to Afshartus (*op. cit.*), 320 groups are needed to estimate reliably the variance components.

The ethnic composition of pupil population was 85.6% (1480) of Greek pupils while the remaining 14.4% (248) was of foreign/repatriated pupils.

‘Decisions about samples - both sample size and sampling strategies depend on prior decisions about the appropriate unit of analysis to study’ (Quinn Patton, 1987, p. 50).

Due to the limited implementation of reception classes in schools during 1999-2000, all schools with reception classes at the beginning of the school year were included in the sample. From the total of 55 schools included in the current sample, in 12 schools a coach class was operating, while in 13 schools a reception class was operating. In four schools from the above, both support schemes were operating. The remaining 28 schools did not provide any form of support teaching to foreign/repatriated pupils.

Some of them provided only support with homework sections and were operating in the framework of the whole-day school. These sections cater for the needs of all pupils, whose parents are working and do not target categorically foreign/repatriated pupils with limited competence in Greek.

2. 9. Response rates

Among the 1858 pupils who participated in the study 1506 pupils (81%) were tested on both occasions in language and 1507 (81%) were tested on both occasions in mathematics. The remaining pupils were absent on one testing occasion, either at the beginning or at the end of the school year.

Questionnaires were administered to every principal. However, 44 from the total of 58 principals provided data in relation to the implementation of support schemes. Hence the response rate is 75%. In relation to the number of schools that were measured on both testing occasions (53) the response rate is 83%. The response rate for support class teachers was 95%.

2. 10. Numbers of foreign and repatriated pupils

The number of foreign and repatriated pupils enrolled in each school varied greatly between schools.

Table 2. 3: Minimum, maximum and mean number of foreign and repatriated pupils in the schools that took part in the study¹⁶.

	<i>Minimum</i>	<i>Maximum</i>	<i>Mean</i>
Foreign pupils:	2	54	22
Repatriated pupils:	0	47	6
<i>Total:</i>	4	71	28

47 principals replied to this question.

The total number of foreign/repatriated pupils enrolled in all grades in a given school may be employed as an indicator representing the degree to which there might be a need for support teaching provision in any given school. This number is merely indicative since some of these foreign/repatriated pupils may already have reached an advanced level in Greek language or may not have accumulated learning gaps in other lessons and therefore might not need support instruction.

¹⁶ Numbers of foreign and repatriated pupils enrolled in all grades of primary schools were introduced in the above table.

2. 11. Confidentiality issues

It is acknowledged that comparisons between individual schools and/or between classrooms are unacceptable from a statistical point of view (Goldstein & al., 1993). Instead comparisons of schools/classrooms with the average are legitimate. For confidentiality reasons schools/ classrooms should be informed only about their position in relation to the average and they should not have access to the relative performance of other schools / classrooms.

Schools in which classrooms performed at a below-average level and schools in which classrooms performed at an average level have been made anonymous in the 2nd appendix. Only schools in which classrooms performed at an above-average level are shown.

2. 12. The need to employ disaggregated (pupil level) data

Educational research has often employed the impact of aggregated variables at the teacher-level or at the school level on pupils' learning outcomes through single level regression models. However, this design poses several threats to statistical validity including: aggregation bias, undetected heterogeneity of regression among sub-units, wrongly estimated parameter estimates and their statistical errors and related problems associated with the failure to satisfy the assumptions of independence required by single-level models (Hill and Rowe, 1996). In addition, if the aim of a study is to assess the impact of individual or pupil factors¹⁷ on the dependent variable and aggregated data is employed instead of pupil-level data, a 'shift of meaning' takes place, which is known as 'ecological fallacy'. Relationships identified between level-1 variables (individual pupils' characteristics and attainment levels) cannot be generalised to express relationships between variables at level-2. Furthermore, any relationship identified applies to the macro-units only and it does not apply to the micro-units. According to Robinson (1950, cited in Snijders et al., 1999), '*a correlation between macro-level variables cannot be used to make assertions about micro-level relation*'. This third type of error is the '*neglect of the original data-structure*' (Snijders, *op. cit.*). Often, relationships between characteristics at level-1 and the dependent variable (in this case pupils' final attainment score) on one hand and between aggregated level-2 characteristics and the dependent variable on the other, point in opposite directions. '*When aggregated data is used, a variable that is aggregated to the macro-level refers to the macro – units, not directly to the micro – units*' (Snijders et al., *op. cit.*, p. 13).

'There is now fairly general acceptance that studies of school effectiveness in cognitive areas require adequate control for prior attainment at the level of the individual student' (Creemers, 1994a, 1994b; McPherson, 1992; Reynolds and Cuttance, 1992; Scheerens, 1992). The current study was based on pupil level data since curriculum-based tests were administered to all the pupils enrolled in the 6th grade on both testing occasions, practice that enabled the researcher to describe relationships between variables at the pupil level in addition to the classroom and to the school level.

17 Which refer to the lower level, in the hierarchical structure pupils within classrooms within schools.

2. 13. Testing for school or classroom effects

Hill and Rowe, (1996) explain that in many School Effectiveness Studies most of the school level variation that turns out to be significant when the total unexplained variation is partitioned into two levels (school level and pupil level), can in fact be attributed to variation between individual classrooms when three levels are employed (school, classroom and pupil level). *'The unique variance due to the influence of the school and not to the classroom, shrinks to very small levels'* (Hill and Rowe, *op. cit.*).

School level factors set up the framework and the conditions facilitating or inhibiting classroom operation (instruction, classroom climate, discipline, etc.).

Good and Brophy (1986) noted that *'researchers need more carefully to isolate school effects from the effects of other levels of schooling (such as the teacher or district), noting that 'the effects of all these on achievement are confounded in reality'* (p. 590).

Sammons (1996) professed that *'It is necessary to include schools with two or more classes per grade to allow separation of variance at the student, class and school levels'*.

For these reasons the simultaneous existence of school or classroom effects is tested in hierarchical models in which the total variability is divided between the school level, the classroom level and the pupil level.

Therefore in addition to school effects, any school effectiveness study should also test whether classroom effects exist. This analytic path has been pursued in the current study.

2. 14. Contextual characteristics of the classrooms

Contextual characteristics, derived from aggregating individual pupils' characteristics for each classroom will be employed as classroom-level variables. Additional contextual characteristics at the classroom level will also be employed. For all classrooms that participated in the study the following contextual characteristics were calculated: pupils' average initial attainment score (expressed in normalised scores), variance between these initial attainment scores, the percentage of foreign/repatriated pupils, the number of foreign/repatriated pupils, the percentage of low achievers, the percentage of high achievers, the number of high achievers, the average weekly amount of homework undertaken, the average rate of absenteeism (average number of absences), the gender composition (percentage of boys), the average educational level of the pupils' mothers, the percentage of pupils from lower social class, and finally the class size.

2. 15. Duration of the study

A school effectiveness study is conducted according to a longitudinal design. The current study was carried out during one school year period (nine months) in order to minimise problems of attrition of pupils enrolled in the schools sampled. The study was conducted during the school year 1999-2000.

2. 16. Time – frame

For the administration of performance-based tests two teaching hours were employed for each subject (language and mathematics). In total four hours at the beginning and four hours at the end were required for each testing occasion in each classroom.

2. 17. Training of research assistants

Two research assistants were employed to assist the researcher in testing the 83 classrooms investigated. The researcher and the assistants administered the performance-based tests, answered pupils' questions and supervised the pupils during each testing occasion. The researcher trained the research assistants in administering the tests, in order to homogenise the way the tests were administered throughout the sample and ensure reliability. Another aim was to ensure that pupils' testing would fit within the two teaching hours that were allocated by each school.

2. 18. Choice of Region

The study was based on the city of Piraeus, which is the port of Athens. Piraeus and Athens, the Greek capital, belong to the prefecture of Attiki.

This choice seemed to be manageable and feasible for one researcher. If the researcher had undertaken a nationally representative study, it would be difficult to collect data in due time. Another difficulty would have been the fact that the performance-based tests would have had to be administered by the teachers themselves. In that case, 'teaching to the test' scenarios would have skewed classroom final attainment and/or progress residual estimates. In addition, there would have been little uniformity in the administration of the tests across schools and questions of reduced reliability would have cropped up. Another limitation associated with the use of mailed questionnaires is that response rates would have been much lower.

Choosing a sample from one city instead of the whole country implies that the study has reduced generalisability and increased reliability. *'Generalisation is an act of reasoning from the observed to the unobserved, from a specific instance to all instances believed to be like the instance in question'* (Schwandt, 1997, p. 105). The aim of this study is that it is generalisable only within the city in which the study was conducted. The city of Piraeus is comprised of only one school district, which serves mainly urban and working class areas. This administrative district includes five education offices, four of which are located on the mainland while one is comprised of the five main islands of the Saronic gulf, namely Salamis, Aegina, Poros, Hydra and Spetses. Schools belonging

to the latter office (in the islands) were not included in the study for reasons of inaccessibility. The remaining area of Piraeus also includes four educational offices, under one district's auspices. The concentration of ethnic minorities in the city of Piraeus is not as high as in other areas of Attiki, such as in the North of Athens. It was estimated that there were three foreign/repatriated pupils in each classroom in the participating schools.

2. 19. Choice of the 6th grade

The 6th grade was chosen because it marks the transition point between primary and secondary education. If foreign/repatriated pupils do not manage to catch up while they are enrolled in primary school, then the gap between their performance and the performance of their Greek colleagues is likely to widen during secondary school, as support teaching programmes targeting them categorically do not exist (Korilaki, 1997). Also, according to data derived from a census survey undertaken by NSSG during the school year 1994-95, which was employed in the previous study, the percentage of Albanian students and students from the former USSR in relation to Greek pupils is smaller in lower secondary education¹⁸ than in primary. The lower participation of foreign/repatriated pupils in secondary education might indicate that they drop out of school at some point after their transition from primary to secondary school. In Greece, compulsory education is comprised of nine grades.

¹⁸ Lower secondary education is comprised of grades 7-9.

2. 20. Pilot study

The performance-based tests employed in the main-study in 1999-2000 were piloted and subsequently revised during the previous year (1998-1999) so that their psychometric properties (difficulty level, discrimination index, etc.) closely match the outcomes obtained from 6th grade pupils. In total, six tests were devised: (four tests measuring pupils' performance in mathematics and two tests measuring pupils' performance in language). From the four tests on mathematics, two were parallel versions of the 5th grade curriculum (those used at the beginning) while the other two were parallel versions of the 6th grade curriculum (those used at the end). The following table summarises the role, purposes of instruments involved as well as the different analyses undertaken in the pilot and in the main study.

Table 2. 4: Role, purposes, instruments involved and analyses undertaken in the pilot and the main study

	Pilot	Main Study
Purpose:	<ul style="list-style-type: none"> To create performance-based tests measuring pupils' attainment in language and mathematics. 	<ul style="list-style-type: none"> To measure differences in attainment and progress between groups of pupils and between classrooms or schools. To look into the implementation of coach classes and reception classes.
Sample:	20 primary schools.	54 primary schools with 83 classrooms.
Instruments:	<ul style="list-style-type: none"> Two performance-based tests for mathematics at the end of year 5 and two tests at the end of year 6. One performance-based test for language at the end of year 5 and one test at the end of year 6. 	<ul style="list-style-type: none"> Two performance-based tests for mathematics at the end of year 5 and two tests at the end of year 6. One performance-based test for language at the end of year 5 and one test at the end of year 6. A principal's questionnaire
Outcomes measured:	Pupils' attainment at the beginning and at the end of the final year (year 6) of primary school in language and in mathematics.	Pupils' attainment at the beginning and at the end of the final year (year 6) of primary school in language and in mathematics.
Period of data collection:	At the beginning and at the end of the 6 th grade	<ul style="list-style-type: none"> At the beginning and at the end of the 6th grade for the administration of the performance-based tests. At the end of the school year for principals' questionnaire.
Analyses undertaken:	Item analysis, reliability indexes etc.	Hierarchical linear models assisted by MLWin software and SPSS.

2. 21. Methodological Summary

To summarise, a longitudinal school effectiveness design has been employed in order to disentangle the impact of pupils attending schools and classrooms from the impact of certain pupils' characteristics (such as ethnicity, social class and gender, nursery school attendance, mother's education, years in Greek school, amount of homework undertaken and absences) on pupils' attainment and progress rates. The attainment of pupils enrolled in the 6th grade was measured twice in language and mathematics, at the beginning and at the end of the school year 1999-2000. A longitudinal design requires that multilevel modelling techniques be employed in the analysis.

3. MULTILEVEL ANALYSIS OF PRIMARY SCHOOL EFFECTIVENESS IN MATHEMATICS AND IN LANGUAGE

3. 1. Issues that will be covered in this analysis

This quantitative analysis closely adheres to school effectiveness methodology. It aims to establish whether school and/or classroom effects exist according to a range of models and to estimate differences in terms of pupils' attainment, adjusted attainment, progress and adjusted progress, for different pupil groups. The analysis will identify the effects that belonging to particular classrooms or schools have on pupils' attainment and on pupils' progress rates during the final year (6th grade) of primary school. The analysis will estimate the percentage of the total variability that can be attributed to the school, the classroom and the pupil level according to a range of models for language and for mathematics; this will be done in order to assess the extent to which schools or classrooms perform similarly in the two subjects as well as the extent to which they perform similarly according to a range of models for any single subject. Furthermore, the analysis of schools or classroom residuals will identify the extent to which schools and classrooms have performed similarly in the two subjects according to a similarly specified model and the extent to which classrooms have performed similarly in the same subject according to a variety of models. Analysis of random effects will reveal whether schools and classrooms are differentially effective for different pupil groups. Predictions according to a variety of models are intended to corroborate the finding of classroom differential effects for different pupil groups and make it empirically salient. Predictions are also intended to investigate empirically whether school effects exist. Corre-

lations between school/classroom residuals derived from different models and correlations between residuals derived from similarly specified models across subjects explore the issue of effectiveness.

The impact of several individual pupils' characteristics on pupils' final attainment scores or on pupils' progress has been estimated individually and jointly and interactions between the main effects terms have been tested in a variety of attainment and progress models. More specifically, the variables tested are ethnicity, social class, highest educational level reached by a pupil's mother, gender, hours of weekly homework undertaken by a pupil, days that each pupil has been absent during the school year, the number of friends a pupil has within the class and whether a pupil has attended nursery school or not. Whereas the first four variables characterise each child as soon as s/he is enrolled in the primary school, the latter four characteristics are related to pupils' previous schooling experience.

3. 2. Multi – level modelling

Multi- level models account for the fact that there is some form of dependency between observations clustered within higher-level units. For example, pupils are clustered within classrooms and within schools; patients are clustered within hospitals, etc.

'The technique of multi-level modelling accounts for this dependency by partitioning the total variance in the data into variation due to these sources or 'higher level units' and the residual variation that remains' (Goldstein, 2001, p. 2). 'The advantages of multilevel modelling are not only that it capitalises on the hierarchical structure of the data but also that this technique can be used to look at potentially interesting differ-

ences, such as those between the performance of males and females having taken account of their attainment on entry, thus allowing a fair comparison of like with like' (Thomas and Mortimore, 1996, p.8).

In the current study the educational attainment of 6th grade pupils will be considered in mathematics and in language. Pupils' attainment is expressed in grades they obtained in tests administered at the beginning and at the end of the school year¹⁹. These grades were subsequently transformed into normal score²⁰ units. Initially, a two-level model will be fitted; that is a model including only the pupil level and the classroom level. In this model the total variation will be divided between variation corresponding to the pupil level and variation corresponding to the classroom level. A three-level model dividing the total variation between the school level, the classroom level and the pupil level could not fit²¹.

19 As also discussed in the methodology chapter, typical reliability coefficients have been obtained for all tests employed. The properties of the tests, the pupils' attainment distributions associated with these tests along with the reliability coefficients obtained are shown in appendix 1.

20 Normal scores have a mean of 0 and a standard deviation of 1 and they have a normal distribution.

21 Full details of all the modelling undertaken will be available to others on request.

3. 3. Examining the distributions of mathematics tests

The total initial and final attainment score distributions derived from mathematics tests seemed to be more symmetrical than their respective distributions derived from language tests. The distributions of the mathematics tests were more centred and they seemed to approach the normal distribution, whereas the distributions of language test scores were negatively skewed. There was a ceiling effect in the attainment distribution of both pupils' initial attainment and in their final attainment scores in the case of language. This ceiling effect was derived from the fact that many items of low and average difficulty, and very few difficult items, were included in both the above language tests. The inclusion of a majority of relatively easy and easy items resulted in average and high scores for the vast majority of pupils tested. However, the fact that the distributions of both language test scores referring to the beginning and the end of the school year are negatively skewed means that the two sets of scores are comparable. In order to correct these deficiencies, inherent in the construction of language tests and to a less extent of mathematics tests and to enable comparisons of the impact of certain variables across subjects, pupils' initial and final attainment scores were transformed to normalised scores. Pupils' final attainment score expressed in normal score units will be used as the dependent variable in all subsequent analyses, whereas pupils' initial attainment score expressed in normal score units will be used as the major explanatory variable in all progress models.

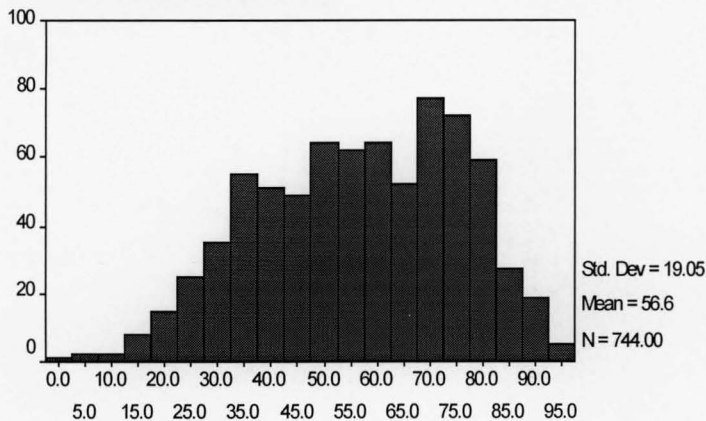
As mentioned in the methodology chapter, the reliability coefficient of the mathematics test scores measuring pupils' performance at the beginning of the school - year was es-

estimated to be 0.88 and 0.92 for parts A and B of tests respectively. The reliability coefficient for mathematics test scores measuring pupils' performance at the end of the school year was estimated to be 0.89 for both parts A and B. The following four histograms show the distributions of the mathematics test scores referring to the beginning and end of the school year and the next two show the distributions of the initial and final attainment test scores for language. All pupils enrolled in the 6th grade were tested.

Total score of A part of mathematics test

Beginning of the school - year

All pupils are included

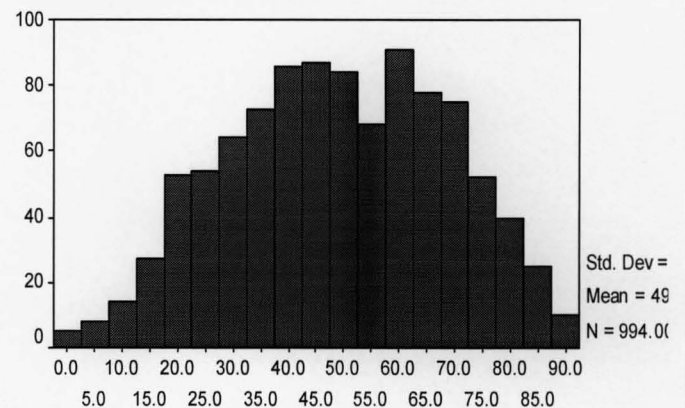


T_SCORE

Total score of B part of mathematics test

Beginning of the school - year

All pupils are included

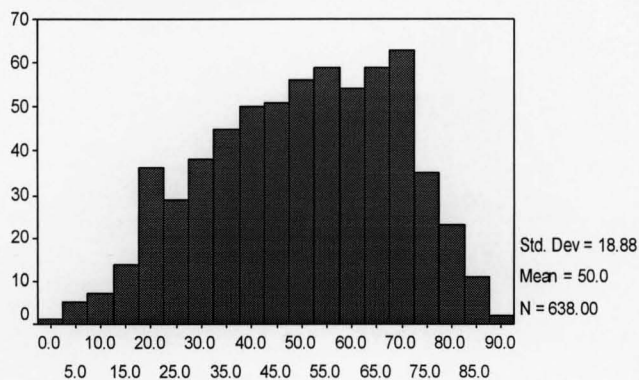


Total score (overall)

Total score of A part of mathematics test

End of the school - year

All pupils are included

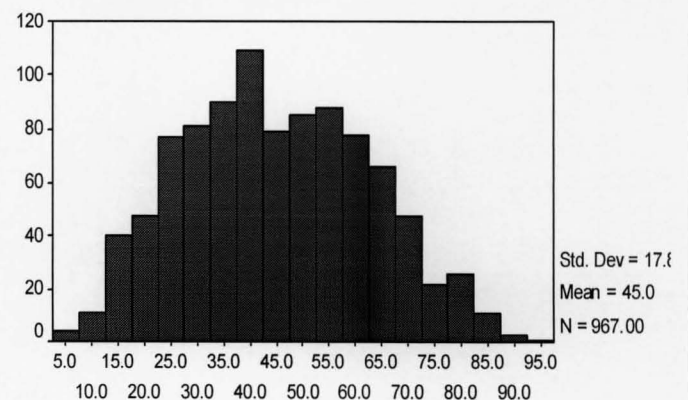


T_SCORE

Total score of B part of mathematics test

End of the school - year

All pupils are included



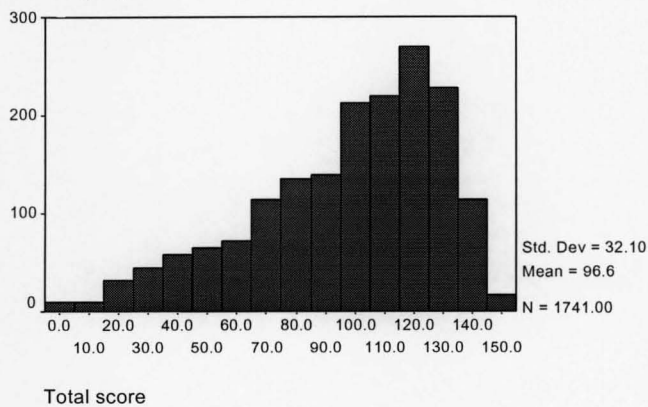
T_SCORE

3. 4. Examining the distributions of language tests

The distributions of language tests were negatively skewed, as these tests comprised a majority of easy or relatively easy items. The reliability coefficient of the language test referring to the beginning of the school year was estimated to be 0.94 whereas the reliability coefficient of the language test referring to the end of the school year was estimated to be 0.93.

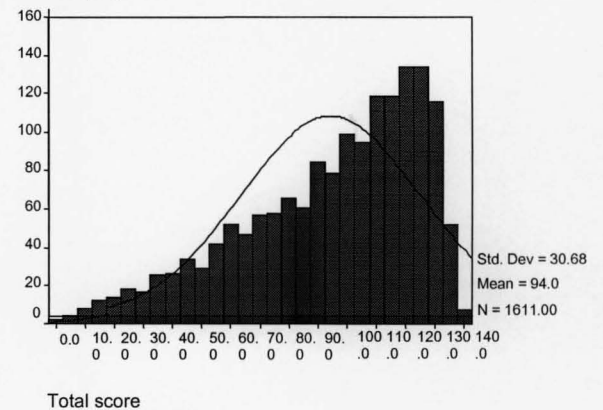
Language total score, Beginning
of the school - year

All pupils



Language total score, End of the school yr.

All pupils



3. 5. Descriptive Statistics of all variables employed in the analysis

Table 3. 1: Mean and standard deviations of variables measuring individual pupils' characteristics

	N	Mean	Std. Deviation
Gender (1: boy, 0: girl)	1758	0.52	0.500
Mother's educational level Level she has successfully completed: (0: no grades – illiterate, 1: a few grades, 2: primary school, 3: compulsory education (9 years) 4: secondary education (12 years) 5: tertiary education, 6: University degree)	1408	3.81	1.212
Nursery School Attendance (1: attended nursery, 0: did not attend nursery)	1578	0.85	0.355
Friends in the class	1738	4.18	2.869
Homework	1525	12.3	9.2
Days of Absence	1264	4.5	5.8
Ethnicity (1: foreign, 0: Greek)	1723	0.11	0.313

The previous table provides descriptive statistics for all pupils.

Table 3. 2: Representation of Greek, Albanian, pupils coming from the democracies of the former USSR, as well as from ‘Other’ ethnic backgrounds, grouped together in the same category, assessed in mathematics at the end of the school-year.

	Frequency	Valid Percent
Greek	1480	86
Albanian	152	8.8
The former USSR	56	3.2
Other ethnic groups	33	1.9
Total	1721	100
Missing values:	137	
Total	1858	

Table 3. 3: Frequencies of pupils’ social class categories according to the 9-category Goldthorpe social class schema. Each pupil is allocated to the higher social class between the two social class positions of his/her parents

Goldthorpe’s social class designation	Description	Frequencies	Valid percent
1	High grade professionals, administrators and officials; managers in large industrial establishments; large proprietors	43	2.7
2	Lower grade professionals, administrators and officials; higher grade technicians; managers in small industrial establishments; supervisors of non-manual employees	148	9.4
3	Routine non manual employees, higher grade (administration and commerce)	163	10.3
4	Routine non manual employees, lower grade (sales and services)	66	4.2
5	Small proprietors, artisans, etc. with employees	172	10.9
6	Small proprietors, artisans, etc. without employees	355	22.4
7	Lower grade technicians; supervisors of manual workers	80	5.1
8	Skilled manual workers	349	22.1
9	Semi- and unskilled manual workers (not in agriculture, etc.)	202	12.8
10	Unemployed	4	.3
	Total:	1582	100.0
	Missing:	276	
	Total	1858	

The median coincided with the 6th category. The 1st percentile coincided with the 4th category, the 2nd with the 6th category and the 3rd with the 8th category.

Table 3. 4: Gender by dichotomous social class for Greek pupils: case processing summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Gender by dichotomous social class	1275	86.1%	205	13.9%	1480	100.0%

Table 3. 5: Cross-tabulation of gender by dichotomous social class

		Social class		
		6,7,8,9,10 manual	1,2,3,4,5 non manual	Total
Female	Count	386	244	630
	% within Female	61.3%	38.7%	100.0%
Male	Count	386	259	645
	% within Male	59.8%	40.2%	100.0%
Total	Count	772	503	1275
	% within Total	60.5%	39.5%	100.0%

Table 3. 6: Cross-tabulation of ethnic group membership by gender

Ethnic groups		GENDER		Total
		FEMALE	MALE	
'Albanians'	Count	63	89	152
	% within Albanian	41.4%	58.6%	100.0%
'Former-USSR'	Count	24	32	56
	% within the former USSR	42.9%	57.1%	100.0%
'Other'	Count	8	25	33
	% within Other	24.2%	75.8%	100.0%
Missing	Count	2	7	9
Total	Count	97	153	250
	% within total	38.8%	61.2%	100.0%

Table 3. 7: Cross-tabulation of Ethnic group by dichotomous social class (1: manual, 0: non-manual)

Ethnic group		Social class		Total
		1,2,3,4,5 non manual	6, 7, 8, 9 manual	
Albanian	Count	22	87	109
	% within Albanian	20.2%	79.8%	100.0%
Former USSR	Count	17	31	48
	% within the former USSR	35.4%	64.6%	100.0%
Other	Count	15	9	24
	% within 'Other'	62.5%	37.5%	100.0%
Missing		3	3	6
Total	Count	57	130	187
	% Total	30.5%	69.5%	100.0%

Table 3. 8. Pearson's Correlations between pairs of variables employed in the analysis

Variables		Ethnic status	Gender	Years in Greek school (including Nursery)	Nursury School Attendance	Mother's Educational level	Days of absence	Friends (Greek & foreign)	Social Class (1: lower, 0: higher)
Ethnic status (1: foreign, 0: Greek)	Correlations:	1.000	.074	-.705	-.236	.009	.097	-.211	.109
	Sig.	.	.002	.000	.000	.740	.001	.000	.000
	N	1728	1728	1726	1570	1309	1259	1728	1728
Gender: (1: Boy, 0: Girl)	Correlations:	.074	1.000	-.036	-.064	-.003	.000	.013	-.002
	Sig.	.002	.	.133	.011	.900	.997	.576	.928
	N	1728	1737	1734	1578	1316	1264	1737	1737
Years in Greek school (including nursery) (From 1 to 7)	Correlations:	-.705	-.036	1.000	.475	.073	-.074	.228	-.094
	Sig.	.000	.133	.	.000	.006	.008	.000	.000
	N	1726	1734	1854	1577	1406	1263	1734	1854
Nursury School Attendance (1: attended nursery, 0: did not attend nursery)	Correlations:	-.236	-.064	.475	1.000	.008	-.016	.123	.002
	Sig.	.000	.011	.000	.	.781	.562	.000	.944
	N	1570	1578	1577	1578	1195	1262	1578	1578
Mother's educational level: (from 0 to 6)	Correlations:	.009	-.003	.073	.008	1.000	.001	.179	.006
	Sig.	.740	.900	.006	.781	.	.965	.000	.822
	N	1309	1316	1406	1195	1408	974	1316	1408
Days of absence	Correlations:	.097	.000	-.074	-.016	.001	1.000	-.125	.045
	Sig.	.001	.997	.008	.562	.965	.	.000	.111
	N	1259	1264	1263	1262	974	1264	1264	1264
Friends (Greek & foreign - from sociogram)	Correlations:	-.211	.013	.228	.123	.179	-.125	1.000	-.085
	Sig.	.000	.576	.000	.000	.000	.000	.	.000
	N	1728	1737	1734	1578	1316	1264	1738	1738
Social Class (1: lower, 0: higher)	Correlations:	.109	-.002	-.094	.002	.006	.045	-.085	1.000
	Sig.	.000	.928	.000	.944	.822	.111	.000	.
	N	1728	1737	1854	1578	1408	1264	1738	1858

3. 6. Types of Models Employed in the Analysis

In order to investigate whether there are attainment differences or differences in progress rates between pupil groups defined by ethnicity, social class and gender and to identify the influence of the classroom or the school attended on pupils' educational outcomes, eight model types were developed. Each of these models sheds light on a different aspect of pupils' educational attainment or of pupils' progress; each model addresses a different question.

- ***Null models*** provide the basis for comparison. With null models the unadjusted pupils' attainment, as well as the average final attainment of classrooms, and/or schools can be estimated without the use of any predictor variables. According to null models unadjusted comparisons between classrooms in terms of pupils' final attainment score can be undertaken.

- ***Absolute attainment models*** determine the impact of a wide range of individual pupils' characteristics on pupils' final attainment score. These models enable also adjusted comparisons between similar classrooms in terms of pupil intake to be undertaken, according to classroom adjusted final attainment scores.

Willms (1992) described such school or classroom differences following such controls as Type A effects. Residuals derived from such models enable the researcher to define how well a pupil with average background is expected to perform in a given school in relation to the performance of similar pupils in other schools. Hutchison (1993) has also named these effects as pupil impact residu-

als. Absolute attainment models are feasible even when a cross-sectional research design is adopted, which investigates the impact of variables measuring individual pupils' characteristics on pupils' final attainment scores at a specific (given) time point. Cross-sectional research involves the measurement of all variable(s) for all cases within a narrow time span so that the measurements may be viewed as contemporaneous. A cross-sectional research design deals with the following questions: a) Are there any differences in the performance of pupil groups at a given time-point? and b) can these differences be alleviated by adjusting for additional variables, that is to say for more individual pupils' characteristics? *'Typically such research design collects information on aspects of social background, gender and occasionally neighbourhood, or social context'* (Drew, 1995). Absolute attainment models are built based on a null model. In absolute attainment models individual pupils' characteristics are introduced as explanatory variables to account for differences in pupils' final attainment scores. The fixed effect coefficients in absolute attainment models estimate the average impact of each of the explanatory variables on pupils' final attainment scores while controlling for the effects of the remaining individual pupils' characteristics, included in the same model.

- ***Progress models*** allow comparisons to be made between classrooms in terms of their pupils' 'learning gain' accomplished within the period investigated. These models can estimate the impact of initial attainment scores on pupils' final attainment scores. These models adjust only for pupils' baseline scores.

- ***Adjusted Progress models*** allow comparisons to be made between classrooms in terms of their pupils' 'adjusted learning gain' controlling for the impact of initial attainment score and additional variables included in the same model. The fixed effect coefficients of explanatory variables representing individual pupils' characteristics estimate the average impact of these explanatory variables on pupils' progress while controlling for the effects of the remaining individual pupils' characteristics, in the same model.

Progress and adjusted progress models adhere to a *longitudinal research design*, which requires that one or several groups of participants be assessed twice, or at several time points.

- ***Contextual models*** measure the impact of classroom contextual characteristics such as the class size, on the progress of all pupils enrolled in the same class during the period investigated. These models enable adjusted comparisons in terms of progress between schools and/or between classrooms serving a similar pupil intake, intake composition and context. Willms (1992) has named school or classroom differences following these additional controls as Type B effects.

SER recognizes the need to take account of differences in pupil background characteristics, as well as pupils' prior attainment (Mortimore et al., 1988, Tizard et al., 1988, Thomas, et al., 1995). But again, one also has to take into account the effects of pupils' clustering in classrooms and schools.

'Multilevel techniques enable the separation and identification of effects of individual pupil characteristics on pupil attainment from school influences' (Goldstein, 1995).

Among the aims of the multilevel analysis was to examine many intermediate models, in order to estimate the joint impact of pupils' individual socio-economic characteristics (ethnicity, social class and gender) on pupils' final attainment scores and on pupils' progress, and to examine how this impact is alleviated in the presence of other variables influenced by schooling.

One reason it is also desirable to assess the joint impact of all explanatory variables at stake (measuring individual pupils' characteristics) on pupils' final attainment scores and on their progress, is the following: In cases where there are high correlations among some of the explanatory variables (where there is multicollinearity), it may be that some of them are better predictors than others. In the current study some variables have significant coefficients in an absolute attainment or in a progress model, when no other variables are included in the model. However, the coefficients of these explanatory variables are substantially reduced and sometimes they became non-significant when more explanatory variables were introduced in these models. This fact does not mean that the former variables do not have an impact on pupils' final attainment scores or pupils' progress.

Collinearity involves the relationship of independent variables (predictors) to one another (Kleinbaum et al., 1998). It stems from high correlation between the independent variables and it is often encountered in regression analysis as well. According to Hair et al. (1995, p. 124) the correlation between the predictor variables included in a regression analysis has an impact on the interpretation of the coefficients derived from this analysis. Multicollinearity makes determining the contribution of each independent variable difficult because the effects of the individual predictor variables are 'mixed' or

‘confounded’ owing to collinearity. High multicollinearity results in higher proportions of shared variance and lower levels of unique variance from which the effects of individual predictor variables can be determined. As the correlation between two independent variables decreases, the collinearity problem becomes less severe. Multicollinearity between school climate items and other variables were identified in many School Effectiveness studies. Brookover et al. (1978, 1979) found that school climate factors were powerful indicators of student achievement. When school climate factors were entered first in regression models, they accounted for 73 per cent of variation at the school level in student achievement. However, school climate accounted for only an additional 4 per cent when entered last in a model already controlling for pupils’ family SES background and for a variable measuring the school’s racial composition. Brookover et al. (*op. cit.*) concluded that much of the variance in school mean achievement attributed to student background variables might instead be due to school environment.

In the current data set, the variables ‘mother’s educational level’ and ‘pupil’s social class’ are correlated ($r = 0.38$). As a result, when social class variable enters into a multilevel model, which already includes mother’s education, the coefficient of the mother’s education variable is reduced, and becomes not significant in a progress model. In fact, number of years in education completed by teachers’ parents and by teachers themselves has been used as a measure of teachers’ socioeconomic status in the Louisiana school effectiveness study by Teddlie and Stringfield (1993). It seems that ‘parents’ highest educational level completed’ measures the same construct as ‘teachers’ social class’ that is to say their socio-economic status. In Van De Jong’s

(1989) school effectiveness study the index of social class status is based on the educational and occupational level of both parents.

The variables 'educational level of the pupil's mother' and 'social class' are highly interrelated. The education and training that a woman receives by virtue of her class background provide a highly significant contribution to the position she will occupy in the labour force (Deem, 1986, p. 175) and hence to her social class. Deem (*op. cit.*), found that the correlation between pupils' mothers' education and pupils' social class was 0.48 while the correlation between pupils' mothers' education and fathers' social class was 0.40. Often social class is measured using parental occupation and mother's and father's terminal age of education (Patterson, 1991b; Smith and Tomlinson, 1989, cited in Hutchison, 1993, p. 31).

Therefore the analyst at a preliminary stage should use simple correlations between the predictors themselves and between each predictor and each dependent variable to understand how other independent variables may mediate this relationship.

3. 7. Null models

Null models 2 (for both mathematics and language):

If y_{ij} is the end score (in mathematics or language) of the i^{th} pupil in the j classroom, then

the following model can be written:

$$y_{ij} = \beta_0 + u_j + e_{ij}$$

The above model can also be expressed by the following equations:

$$y_{ij} = \beta_j + e_{ij}$$

$$\beta_j = \beta_0 + u_j$$

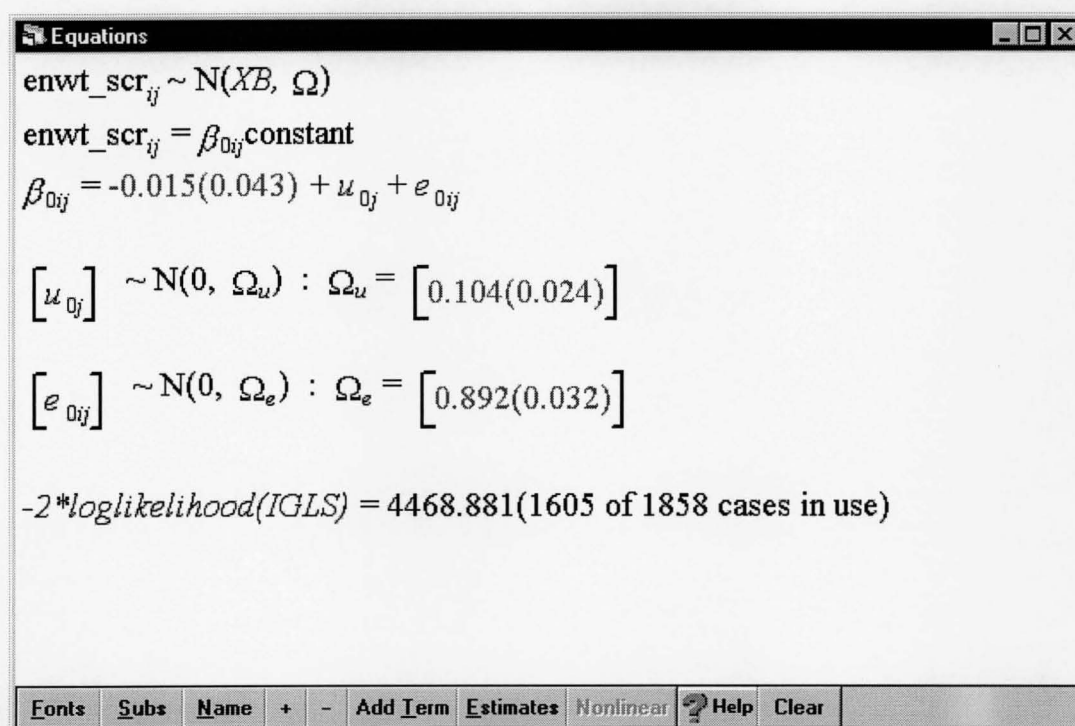
(Goldstein, 1997, p. 378)

To explain the last two equations, the end score consists of the mean of classroom j β_j and a deviation e_{ij} for each pupil from their classroom's mean. The classroom mean β_j can be broken down into the overall classroom mean β_0 and the classroom residual u_j , which is the difference between the overall and the classroom mean.

Classroom-level residuals u_j and pupil level residuals e_{ij} are assumed to be normally distributed.

Figure 1 below shows a window obtained from the MLWIN software, which gives the parameter estimates for the null model obtained through an iterative procedure. The model has converged.

Figure 1: Null model 2 for mathematics



Equations

$$\text{enwt_scr}_{ij} \sim N(XB, \Omega)$$

$$\text{enwt_scr}_{ij} = \beta_{0ij} \text{constant}$$

$$\beta_{0ij} = -0.015(0.043) + u_{0j} + e_{0ij}$$

$$\begin{bmatrix} u_{0j} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} 0.104(0.024) \end{bmatrix}$$

$$\begin{bmatrix} e_{0ij} \end{bmatrix} \sim N(0, \Omega_e) : \Omega_e = \begin{bmatrix} 0.892(0.032) \end{bmatrix}$$

$$-2 * \loglikelihood(IGLS) = 4468.881(1605 \text{ of } 1858 \text{ cases in use})$$

Fonts Subs Name + - Add Term Estimates Nonlinear ? Help Clear

The first line in the main body of the window specifies the default distributional assumption: the response vector has a mean specified in matrix notation by the fixed part XB , and a random part consisting of a set of random variables described by the covariance matrix Ω . This covariance matrix Ω incorporates the separate variances of the random coefficients at each level (Rasbash et al., 2000, p. 20). The second line in the window relates the dependent variable to a set of explanatory variables. In this case, no explanatory variable is included, and the response variable is related to a constant vector of ones, multiplied by the coefficient β_{0jk} . The next (3rd) line informs us that the model is made multilevel by allowing each classroom's summary line to depart (be raised or lowered) from the average line by the amount u_{0j} . The i^{th} pupil in the j^{th} classroom departs from the classroom's summary line by an amount e_{0ij} . Classroom intercepts β_{0k} are formed by adding the classroom level residuals u_{0j} to the constant (-0.015). The 4th, and

5th lines show the variances at the 2nd and 1st level, that is to say the variances at the classroom and at the pupil level.

The last line shows the value of the -2loglikelihood statistic, which is used to compute the deviance test. The deviance test enables the researcher to assess the goodness-of-fit of successive nested models. The difference between two -2loglikelihood statistics (in two models fitted on the same number of cases) is known as the deviance statistic. This deviance statistic allows the goodness of fit of a new model to be estimated in relation to a well-established previous model. The deviance statistic follows a chi-square distribution with degrees of freedom equal to the additional number of parameters that have been introduced in the later model.

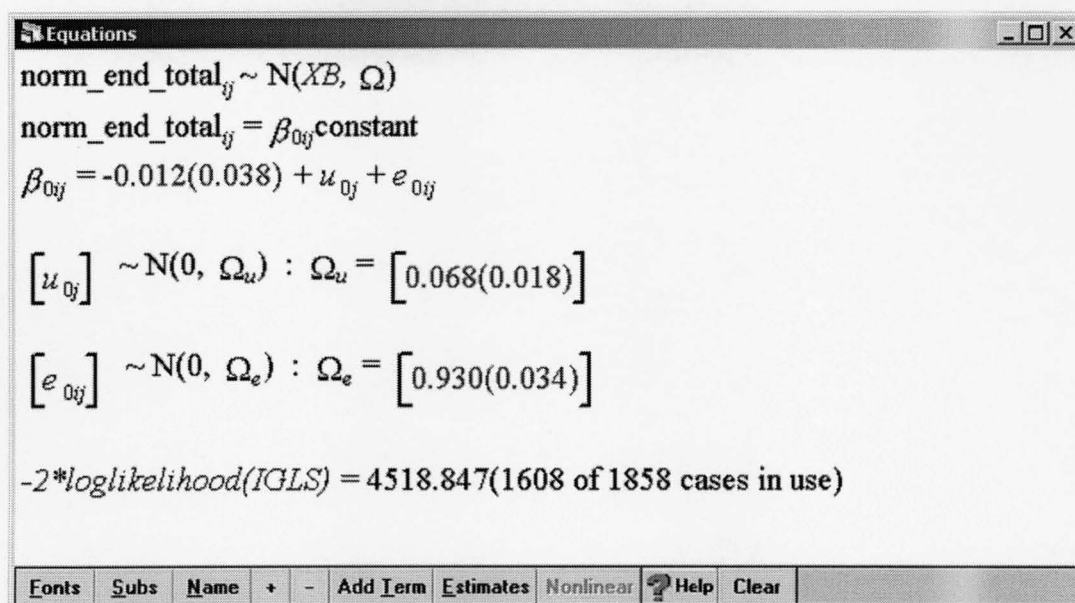
This Null model serves as the starting point; it is the basic empty model that will be used to assess the goodness of fit of following models. Although it does not contain any explanatory variables, the null model is important because it partitions the variability encountered in the dependent variable between the levels considered (in this case at the classroom and at the pupil level). It shows that the classroom level variability is relatively small compared to the pupil level variability.

The 253 missing cases reported are pupils who missed either the beginning or the final testing occasion, as they were absent on one testing day. Their attainment profile differs slightly, but not a lot, from those of pupils who were tested on both occasions. The average initial score of pupils who were tested only at the beginning is -0.256. Pupils who were tested only once are from a variety of ethnic backgrounds, namely 102 (88%) of them Greek and 29 (22%) of them foreign, whereas the actual percentage of foreign/repatriated pupils in the whole sample is 14.4%, while in total 248 for-

eign/repatriated pupils were included in the sample. In that sense, there is a slight over-representation of foreign/repatriated pupils among the pupils who did not sit the final exam at the end of the school year. The remaining 122 were pupils identified to have special educational needs.

Figure 2 below shows the basic empty model in which the total variability has been divided into two levels for language.

Figure 2: Null model 2 for language



Equations

$$\text{norm_end_total}_{ij} \sim N(XB, \Omega)$$

$$\text{norm_end_total}_{ij} = \beta_{0ij} \text{constant}$$

$$\beta_{0ij} = -0.012(0.038) + u_{0j} + e_{0ij}$$

$$[u_{0j}] \sim N(0, \Omega_u) : \Omega_u = [0.068(0.018)]$$

$$[e_{0j}] \sim N(0, \Omega_e) : \Omega_e = [0.930(0.034)]$$

$$-2*\loglikelihood(IGLS) = 4518.847(1608 \text{ of } 1858 \text{ cases in use})$$

Fonts Subs Name + - Add Term Estimates Nonlinear ? Help Clear

In the case of language, the 250 missing cases reported were pupils who missed either the initial or the final testing occasion, as they were absent on one testing day. Their attainment profile does not differ much from those of pupils who were tested on both occasions. The average final attainment score of pupils who were tested on both occasions is 0.022 and their average initial attainment score is 0.041.

These pupils come from all ethnic backgrounds, namely 199 (80%) of them Greek and 38 (16%) of them foreign, whereas the actual percentage of foreign/repatriated pupils in

the whole sample is 14%, and in total 259 foreign pupils were included in the sample.

In that sense there is a slight over-representation of foreign/repatriated pupils among pupils who were tested only once. From the remaining 13, 10 pupils had been identified as having special educational needs while three pupils did not complete the performance-based tests.

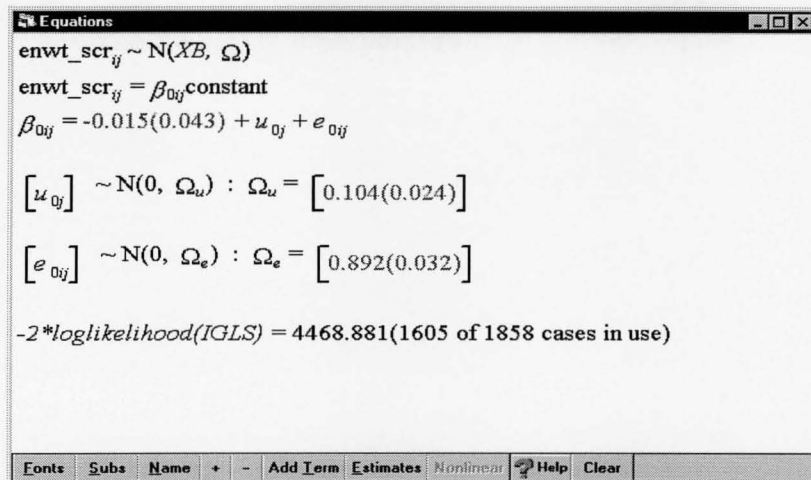
The above null models 2 partition the total variation into two levels: the classroom level (level two) and the pupil level (level one). These null models estimate classroom average final attainment scores. When further explanatory variables are added to subsequent absolute attainment, progress or adjusted progress models, the amount of variability at level two or at level one is reduced, as these additional variables account for some variability. In subsequent phases of the analysis the variability at level two or at level one is unexplained variability that remains after controlling for the effects of explanatory variables.

'The partition of unexplained variability over the various levels is the essence of hierarchical random effect models' (Snijders et al., 1999 p. 46). *'Working with models that do not adequately represent this hierarchical structure 'is dangerous at best, and disastrous at worst'* (Aitkin and Longford, 1986, p. 42, cited in Scheerens and Bosker, 1997, p. 319).

For both subjects Null models 2, which are shown in figures 1 and 2 are improvements on their corresponding Null models 1, in which the total variability has been allocated to the pupil level only. For mathematics, the deviance statistic for the null model 2 in comparison to null model 1 (shown below in figure 3) is significant ($X_1^2 = 75.4$, $p < 0.001$). The same statement holds for the deviance statistic for null model 2 in com-

parison with null model 1 (shown below in figure 4) for language ($X_1^2 = 38.5$, $p < 0.001$). According to these deviance tests null models 2 for each subject are compared with their respective null models 1, in which the total variation has been allocated to the pupil level only. Such a comparison proves that for both subjects null model 2 fits the data better than its respective null model 1.

Figure 1: Null model 2 for mathematics



Equations

$$\text{enwt_scr}_{ij} \sim N(XB, \Omega)$$

$$\text{enwt_scr}_{ij} = \beta_{0ij} \text{constant}$$

$$\beta_{0ij} = -0.015(0.043) + u_{0ij} + e_{0ij}$$

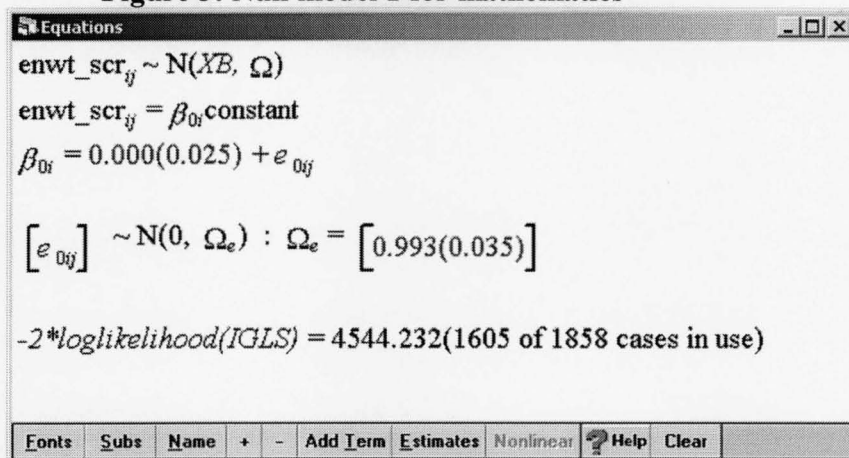
$$[u_{0ij}] \sim N(0, \Omega_u) : \Omega_u = [0.104(0.024)]$$

$$[e_{0ij}] \sim N(0, \Omega_e) : \Omega_e = [0.892(0.032)]$$

$-2 * \log\text{likelihood(IGLS)} = 4468.881(1605 \text{ of } 1858 \text{ cases in use})$

Fonts Subs Name + - Add Term Estimates Nonlinear Help Clear

Figure 3: Null model 1 for mathematics



Equations

$$\text{enwt_scr}_{ij} \sim N(XB, \Omega)$$

$$\text{enwt_scr}_{ij} = \beta_{0i} \text{constant}$$

$$\beta_{0i} = 0.000(0.025) + e_{0ij}$$

$$[e_{0ij}] \sim N(0, \Omega_e) : \Omega_e = [0.993(0.035)]$$

$-2 * \log\text{likelihood(IGLS)} = 4544.232(1605 \text{ of } 1858 \text{ cases in use})$

Fonts Subs Name + - Add Term Estimates Nonlinear Help Clear

Figure 2: Null model 2 for language

Equations

$$\text{norm_end_total}_{ij} \sim N(XB, \Omega)$$

$$\text{norm_end_total}_{ij} = \beta_{0ij} \text{constant}$$

$$\beta_{0ij} = -0.012(0.038) + u_{0ij} + e_{0ij}$$

$$[u_{0ij}] \sim N(0, \Omega_u) : \Omega_u = [0.068(0.018)]$$

$$[e_{0ij}] \sim N(0, \Omega_e) : \Omega_e = [0.930(0.034)]$$

$$-2 * \log\text{likelihood(IGLS)} = 4518.847(1608 \text{ of } 1858 \text{ cases in use})$$

Fonts Subs Name + - Add Term Estimates Nonlinear ? Help Clear

Figure 4: Null model 1 for language

Equations

$$\text{norm_end_total}_{ij} \sim N(XB, \Omega)$$

$$\text{norm_end_total}_{ij} = \beta_{0i} \text{constant}$$

$$\beta_{0i} = 0.000(0.025) + e_{0ij}$$

$$[e_{0ij}] \sim N(0, \Omega_e) : \Omega_e = [0.996(0.035)]$$

$$-2 * \log\text{likelihood(IGLS)} = 4557.315(1608 \text{ of } 1858 \text{ cases in use})$$

Fonts Subs Name + - Add Term Estimates Nonlinear ? Help Clear

A convenient summary of the importance of classrooms is the ratio of the variation at level 2 to the total variation, called the 'Variance Partition Coefficient' (VPC) (Goldstein, 2002, p. 3). The VPC is given by the formula:

$$VPC = \sigma_u^2 / (\sigma_u^2 + \sigma_e^2)$$

Where $\text{var}(u_j) = \sigma_u^2$, $\text{var}(e_{ij}) = \sigma_e^2$

The VPC showing the percentage of total variance attributed to the classroom level for null model 2 for mathematics is:

$$\text{VPC} = 0.104 / (0.104 + 0.892) = 10.4\%.$$

Whereas, for null model 2 for language it is:

$$\text{VPC} = 0.068 / (0.068 + 0.930) = 6.8 \%$$

If the VPC(s) based on null model 2 for mathematics and null model 2 for language are contrasted, it seems that classrooms have a greater impact on pupils' final attainment scores for mathematics than on those for language.

The values of the VPC represent the classroom effect. The above values of VPC are comparable to the results reported by Teddlie and Stringfield (1993, p. 25) that the percentage of the total variation that can be attributed to teachers for the American setting is 12%. The difference is that unlikely to the Greek setting in the American setting school effects were also identified.

In England, Strand (1998, p. 128) reported that at the end of Key Stage 2, according to the null model (in terms of absolute attainment) the VPC, in this case representing percentage of total variance that can be attributed to school membership, was 13.3% for mathematics and 13.2% for language.

In a Dutch longitudinal school effectiveness study, Van Damme et al. (2001) reported on the percentage of total variation that could be attributed to the school and to the

classroom level for language in Flemish secondary schools. According to the 'empty' model the percentage of total variation accounted by school level was 28.3%, by the classroom level was 25.8% and by the student level 45.9%. However, when explanatory variables were introduced in subsequent phases of modelling, most of this variation at both the school and at the classroom level was accounted for by the variables entered. After accounting for student level, classroom level and school-level variables, the remaining unexplained variation at the school level was reduced to 3.8% of the initial variation, whereas the remaining unexplained variation at the classroom level was 5.5% and at the student level was 30.3%.

The above values of VPC for the Greek primary schools' setting are within the expected range for international School Effectiveness studies (for the empty or null model), where 'values between 5% and 20% are common' (Snijders *et al.*, 1999 p. 46). The only difference is that for the majority of other international School Effectiveness studies VPC(s) measure the amount of total variation that can be attributed to schools (school effects), rather than to classrooms (classroom effects).

3. 8. Evaluating the relative standing of classrooms according to progress models with random intercepts

a) Progress Model 1 for mathematics:

Progress models 1 contain unexplained variability at the classroom and the pupil level.

$$y_{ij} = \beta_0 + \beta_1 x_{ij} + u_j + e_{ij}$$

The statistical model in this case becomes:

In which the quantity

$$\beta_0 + u_j$$

is the intercept for the line of each classroom j.

(Goldstein, 1997, p. 383)

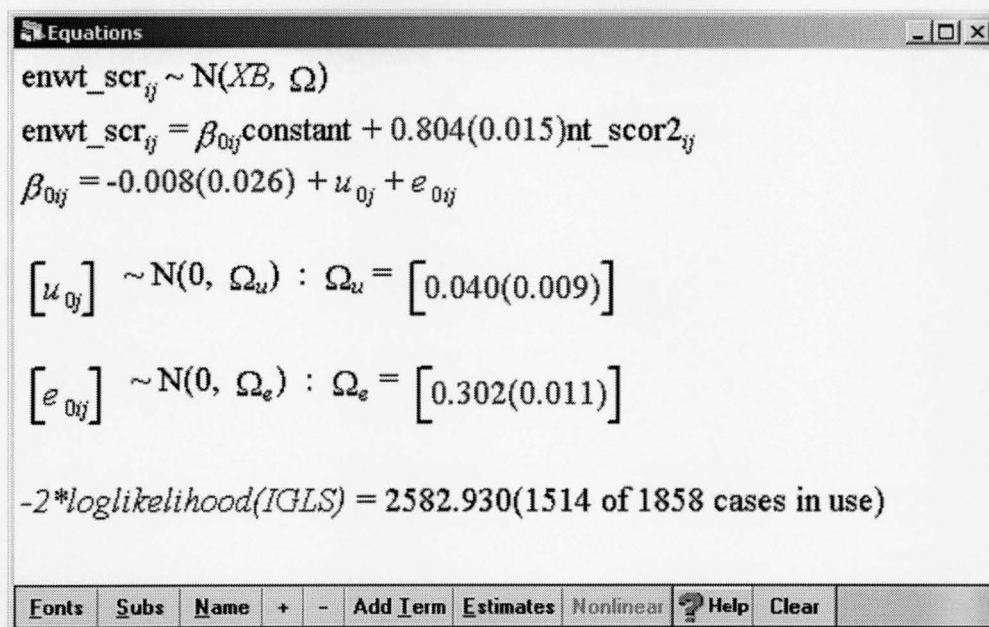
x_{ij} is pupils' initial attainment and β_1 represents the average predicted increase in pupils' final attainment score for a unit increase in their initial attainment score.

According to this model the relationship between pupils' initial attainment score and their final attainment score does not vary between classrooms.

In order to assess the goodness of fit of the above progress models, these models should be compared with their respective null models 2 (in which the total variability has been allocated to two levels), which are based on the same number of cases.

'Progress models 1' shown in figures 5 and 6 below examine whether some classrooms contribute to creating more progress for the pupils enrolled in them than other classrooms.

Figure 5: Progress model 1 for mathematics



The screenshot shows a software window titled "Equations" with the following content:

$$\text{enwt_scr}_{ij} \sim N(XB, \Omega)$$

$$\text{enwt_scr}_{ij} = \beta_{0ij} \text{constant} + 0.804(0.015) \text{nt_scor2}_{ij}$$

$$\beta_{0ij} = -0.008(0.026) + u_{0j} + e_{0ij}$$

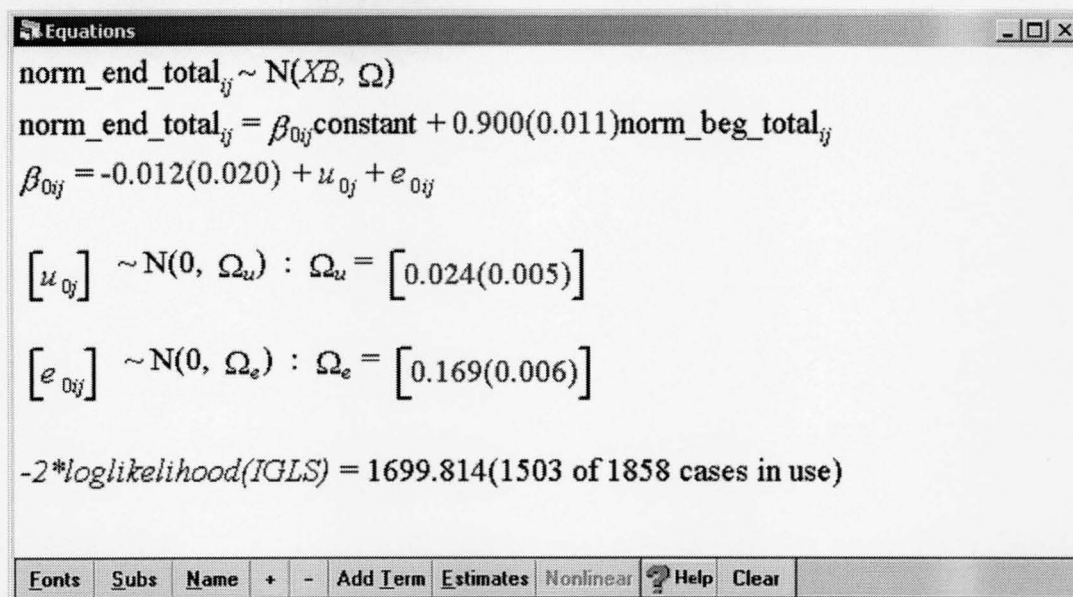
$$[u_{0j}] \sim N(0, \Omega_u) : \Omega_u = [0.040(0.009)]$$

$$[e_{0ij}] \sim N(0, \Omega_e) : \Omega_e = [0.302(0.011)]$$

$$-2 * \loglikelihood(IGLS) = 2582.930(1514 \text{ of } 1858 \text{ cases in use})$$

At the bottom, there is a toolbar with buttons: Fonts, Subs, Name, +, -, Add Term, Estimates, Nonlinear, Help, and Clear.

Figure 6: Progress model 1 for language



The screenshot shows a software window titled "Equations" with the following content:

$$\text{norm_end_total}_{ij} \sim N(XB, \Omega)$$

$$\text{norm_end_total}_{ij} = \beta_{0ij} \text{constant} + 0.900(0.011) \text{norm_beg_total}_{ij}$$

$$\beta_{0ij} = -0.012(0.020) + u_{0j} + e_{0ij}$$

$$[u_{0j}] \sim N(0, \Omega_u) : \Omega_u = [0.024(0.005)]$$

$$[e_{0ij}] \sim N(0, \Omega_e) : \Omega_e = [0.169(0.006)]$$

$$-2 * \loglikelihood(IGLS) = 1699.814(1503 \text{ of } 1858 \text{ cases in use})$$

At the bottom, there is a toolbar with buttons: Fonts, Subs, Name, +, -, Add Term, Estimates, Nonlinear, Help, and Clear.

This time, deviance tests employed tested whether these progress models were improvements on the corresponding null models for mathematics and language. Both the differences in the deviance test between progress model 1 and respective null model 2 are significant at $p < 0.001$, when assessed against a chi-square statistic with 1 degree of freedom ($X_1^2 = 1886$ for mathematics and $X_1^2 = 2475.9$ for language).

Thus ‘progress model 1’ for mathematics and ‘progress model 1’ for language are improvements on the corresponding null model 2 for mathematics and for language at $p < 0.001$. Both the above progress models estimate the impact of pupils’ initial attainment scores on pupils’ final attainment scores. It is generally found that pupils’ *‘initial attainment score is the most powerful predictor of attainment score at the end of a period of schooling’* (Goldstein, 1997, p. 382). According to Willms (1992, p. 58) *‘measures of pupils’ prior performance are essential for statistical control’* if the goal of the analysis is to estimate school (or classroom) effects.

The classroom residuals u_{0j} deriving from the above progress models estimate the difference between the average predicted progress for each classroom controlling only for pupils’ initial attainment score and the average predicted progress for all classrooms.

These residuals u_{0j} can be regarded as group (classroom) effects that are left unexplained by the explanatory variable x_{ij} (pupils’ initial performance).

These residuals, or random errors contain those parts of the variability that are not modelled explicitly as a function of explanatory variables.

Willms (1992, p. 34) stated that: *'A preferable indicator of a school's performance is the distribution of the rates of growth of its pupils, rather than the distribution of pupils' scores on one occasion'*.

L. Stoll and P. Mortimore (1997, p. 9) defined an effective school as *'one in which pupils progress further than might be expected from consideration of its intake'* (in comparison with other schools serving pupils with similar attainment level at entry and socio-economic characteristics). By the same rationale, an effective classroom is a classroom in which pupils performed better than expected based on predictions according to their individual characteristics.

In multilevel modelling the usual practice is to add one explanatory variable at a time thus creating a series of consecutive models and assessing them using the deviance statistic. Later in the analysis, the coefficients of explanatory variables introduced in the fixed part will be allowed to vary randomly at level 2 or at level 1, thus allowing for a set of random parameters to be introduced into the multilevel model.

3. 9. Drawing graphs of predicted classroom lines according to progress models with random intercepts

Based on previously defined progress models 1, shown in figures 5 and 6, a line can be drawn for each classroom showing the relationship between pupils' initial attainment score and their final attainment score.

The progress model presently specified is given by the equation:

$$y_{ij} = \beta_0 + \beta_1 x_{ij} + u_j + e_{ij}$$

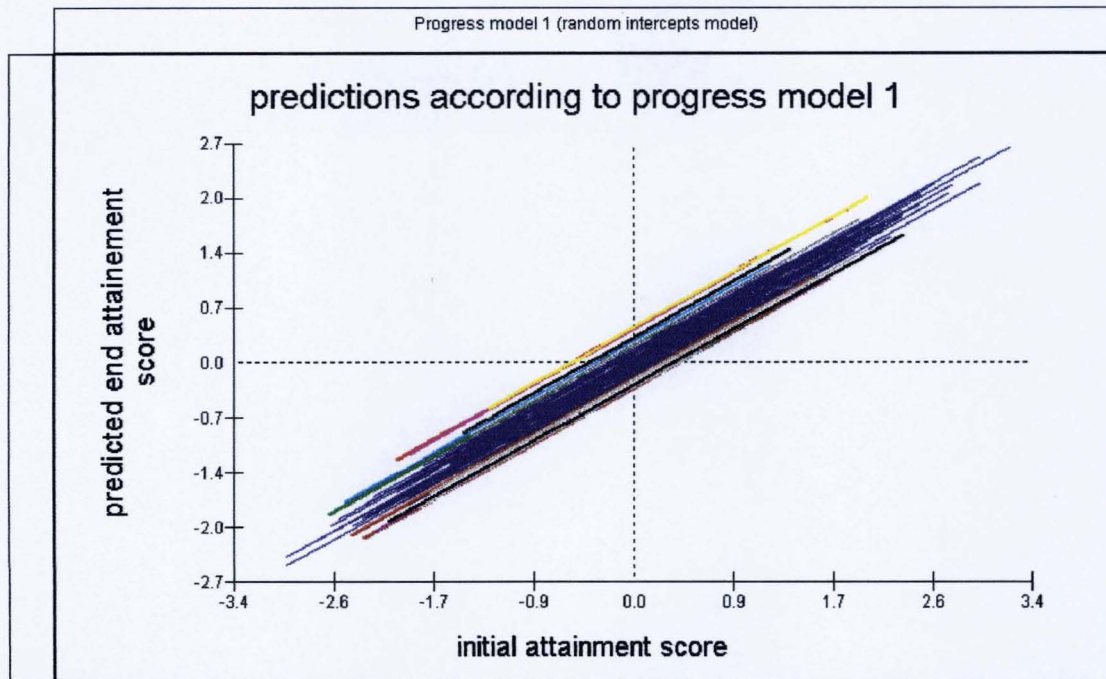
In which the quantity

$$\beta_0 + u_j$$

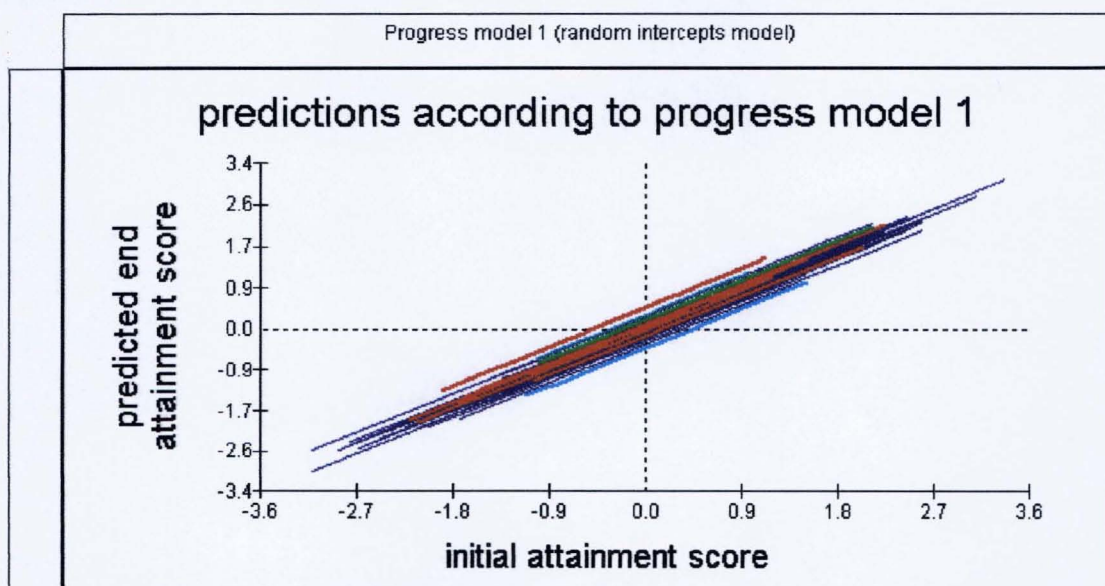
is the intercept for the line of each classroom j

It is evident that classroom lines differ in their intercepts, as the estimated value of different classroom residuals u_j are added to the average estimated intercept β_0 .

Graph 1: End score predictions according to 'progress' model 1, for mathematics.



Graph 2: End score predictions according to progress model 4 for language



In both the above graphs of predicted class lines, classroom lines are parallel; they have the same slopes but they differ in their intercepts. According to the way 'progress mod-

els 1' have been specified, the initial attainment score is not allowed to vary at the classroom level. In other words, the relationship between pupils' final attainment scores and their initial attainment scores is the same for every classroom in the sample. The above are random-intercepts models and they enable comparisons between classrooms in terms of progress, given pupils' initial attainment scores.

For both mathematics and language the coefficients of pupils' initial attainment at the beginning of the school year are rather large (0.804 for mathematics and 0.900 for language). The reduction in deviance in relation to null model 2 is huge (1600.5 at $p < 0.001$ for mathematics and 2475.9, at $p < 0.001$ for language). Thus, both progress models 1 are significant improvements on the corresponding null models 2.

3. 10. Evaluating the relative standing of classrooms according to a progress model with random coefficients (progress model 2) for mathematics

This is a model with initial attainment score as a single explanatory variable in the fixed part as in '*progress model 1*', but with one additional elaboration: The relationship between pupils' initial score and their final attainment score is allowed to vary at the classroom level (between classrooms). Sometimes this model is referred to as the '*random intercepts and slopes model*' or the '*random coefficients*' model or as '*the differential effectiveness*' model.

In SER carried out in the U.K., Nuttall et al. (1990) noted that *‘some schools were more effective in raising the achievement of students with high attainment at entry than that of those with low attainment at entry’* (in Harris, 2001, p. 9).

It was argued that *‘It is not appropriate to talk of the effectiveness of a single school, as though effectiveness was measured on a single dimension and as though the school was equally effective for all groups of pupils. Rather one must investigate the differential effectiveness of schools’* (Nuttall et al., 1990, p. 19, cited in Harris, 2001).

In statistical terms, and if it is considered that up to now in the current Greek data-set classroom effects rather than school effects are found, this model can be described by

$$y_{ij} = \beta_0 + \beta_{1j}x_{ij} + u_{0j} + e_{ij}$$

the following equation (Goldstein, 1997):

In which the intercept and the slope associated with initial score are allowed to vary from classroom to classroom.

The intercept for classroom j , β_{0j} consists of a fixed and a random component:

$$\beta_{0j} = \beta_0 + u_{0j}$$

and similarly the slope of x_{ij} for classroom j , β_{1j} consists of a fixed and a random

$$\beta_{1j} = \beta_1 + u_{1j}$$

component:

Both ‘intercept’ and ‘slope’ terms β_{0j} and β_{1j} , have acquired a subscript j , indicating that they vary according to the classroom with which they are associated, and there is an overall population mean intercept β_0 and an overall mean slope β_1 .

Classroom-level residuals u_{0j} and u_{1j} and pupil-level residuals e_{ij} are normally distributed.

The above model differs from progress model 1 in that there is now a random coefficient associated with pupils’ initial score at the classroom level. Thus classrooms contribute differently to their pupils’ progress in relation to their initial attainment scores. This effect is amply demonstrated in graph 8 where some predicted classroom lines have steeper slopes than others. This signifies that while in some classrooms initially high achieving pupils progress more than average and low achievers, in other classrooms initially low achieving pupils and pupils with average initial attainment progress relatively more.

More specifically one can examine how the relationship between pupils’ final attainment score and their initial attainment score varies between classrooms. This can be achieved by examining the classroom-level variance associated with prior attainment.

Figure 7: Progress model 2 for mathematics

Equations

$$\text{enwt_scr}_{ij} \sim N(XB, \Omega)$$

$$\text{enwt_scr}_{ij} = \beta_{0ij} \text{constant} + \beta_{1j} \text{nt_scor2}_{ij}$$

$$\beta_{0ij} = 0.000(0.027) + u_{0j} + e_{0ij}$$

$$\beta_{1j} = 0.816(0.022) + u_{1j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} 0.044(0.010) & 0.004(0.005) \\ 0.004(0.005) & 0.021(0.006) \end{bmatrix}$$

$$\begin{bmatrix} e_{0ij} \end{bmatrix} \sim N(0, \Omega_e) : \Omega_e = [0.282(0.011)]$$

$-2 * \log \text{likelihood(IGLS)} = 2550.730(1514 \text{ of } 1858 \text{ cases in use})$

Fonts Subs Name + - Add Term Estimates Nonlinear Help Clear

Progress model 2, which is shown in figure 7 above, is a random coefficients (intercepts and slopes) model. The deviance test showed that this model is an improvement on the respective random intercepts model (progress model 1) at $p < 0.001$. The p-value comes from a chi-square distribution with 2 degrees of freedom, as this model employed two extra parameters compared to progress model 1; these parameters are the two additional random coefficients ($X^2_2 = 31.3$).

The coefficients of the fixed part of the model (intercept = 0.000 and slope = 0.816) represent the intercept and the slope of the average classroom. According to this model, an increase of one normal score unit in pupils' initial attainment scores is associated with a gain of 0.816 units in pupils' final attainment scores.

The existence of a random coefficient associated with initial score at the classroom level implies that while some classrooms do better in promoting the attainment of initially high achieving pupils, other classrooms promote more the attainment of initially average or low achieving pupils.

A random coefficient associated with pupils' initial attainment scores at the pupil level did not fit at $p=0.05$ (progress model 2b, $X^2_2=4.7$, $p=0.095$). Therefore, there is little evidence that level-1 residuals are heteroscedastic, in other words that the variance at the pupil level is not constant. Pupil level variance is not a function of initial attainment score, but instead it remains the same (0.282) for all the values of initial attainment score. As will be shown later, this is not the case for language, where pupils' variance will be modelled as a function of initial score. In addition, for mathematics, based on progress model 2, a quadratic term involving pupils' initial attainment scores could not fit ($p=0.075$). Hence, progress model 2 with the effect of initial attainment score random at the classroom level (level 2) is the model that will be used as a base model in subsequent phases of modelling in order to assess the impact of individual pupils' characteristics and contextual effects on pupils' progress. In progress model 2 the classroom-level variance (which is estimated at the average initial attainment score) is reduced by 57.7 % while pupil level variance is reduced by 68.4 % in relation to null model 2.

3. 11. Variance function for mathematics at the classroom level according to progress model 2

In multi-level models the assumption of homoscedasticity encountered in multiple regression, according to which the residual variance is constant, is relaxed and is replaced with the weaker assumption that variances depend linearly or quadratically on explanatory variables. According to progress model 2 the classroom-level (level 2) variance is a quadratic function of pupils' initial attainment score. The following equation demonstrates this relationship:

$$\text{var}(u_{0j}x_0 + u_{1j}x_{1ij}) = \sigma_{u0}^2 x_0^2 + 2\sigma_{u01} x_0 x_{1ij} + \sigma_{u1}^2 x_{1ij}^2$$

Since $x_0 = 1$ the above formula becomes:

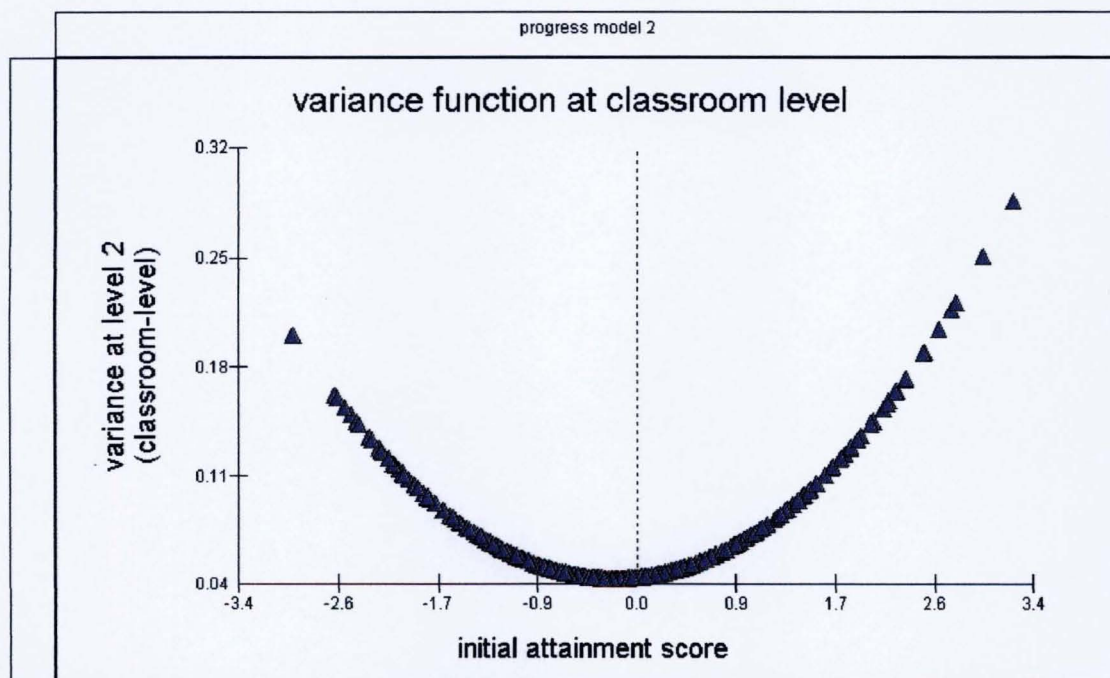
$$\text{var}(u_{0j}x_0 + u_{1j}x_{1ij}) = \sigma_{u0}^2 + 2\sigma_{u01} x_0 x_{1ij} + \sigma_{u1}^2 x_{1ij}^2$$

The existence of random coefficients associated with pupils' initial attainment score at the classroom level implies that the variability among classroom predicted final attainment scores may be modelled through a quadratic function of pupils' initial attainment scores. Graph 8 shows that some of predicted classroom lines have steeper slopes while other classroom lines have less steep slopes. Some classrooms create more progress for

initially highly achieving pupils while in other classrooms the outcomes of initially average and low achieving pupils are promoted more.

No variance components associated at the pupil-level (level 1) were identified.

The diagram below shows that the variance function at the classroom level is a parabola and reaches a minimum of 0.043 at an initial score of approximately -0.2. The variance then increases for values greater than and less than -0.02. Hence, for classrooms in which pupils performed either at a high or at a low level in terms of pupils' initial attainment scores, there is more uncertainty about final attainment scores than for classrooms in which pupils initially performed at an average level.



Graph 3: Level-2 variance function by initial attainment score according to progress model 2 for mathematics

3. 12. Variance at the pupil level for mathematics

According to progress model 2 the variance at the pupil level is constant. It is estimated to take the value $\sigma_{e0}^2=0.282$ for all the values of initial score.

If a graph is created depicting the level 1 variance against pupils' initial attainment scores then a line parallel to the horizontal axis of initial attainment score would emerge cutting the vertical axis, representing the variance at level 1 at 0.282.

3. 13. Variance Partition Coefficient

The VPC is a function of the variance since it reflects the partitioning of unexplained variation at level 2 to the total variation. Since the level 2 in the present study represents the classroom level (only classroom effects can be identified), the VPC expresses the percentage of total variation that can be attributed to the classroom level.

Given that *for mathematics* the classroom-level variance is a function of pupils' initial attainment scores, the VPC is also a function of pupils' initial attainment scores.

Goldstein (2002, p. 5) showed that, while the VPC coincides with the intra class correlation coefficient (ICC) for random intercept models, this is not the case for random coefficient models. In the former case the VPC provides a summary of the importance of classroom variation in relation to the overall variation while at the same time it is also an estimate of the residual correlation between the responses from two students in the same classroom, known as 'intra-unit correlation'.

In the latter case, (e.g. for progress model 2 for mathematics and progress model 4 for language) the intra-unit correlation is given by:

$$\frac{(\sigma_{u0}^2 + \sigma_{u01}(x_{li_1j} + x_{li_2j}) + \sigma_{u1}^2 x_{li_1j} x_{li_2j})}{((\sigma_{u0}^2 + 2\sigma_{u01}x_{li_1j} + \sigma_{u1}^2 x_{li_1j}^2 + \sigma_{e0}^2) \cdot (\sigma_{u0}^2 + 2\sigma_{u01}x_{li_2j} + \sigma_{u1}^2 x_{li_2j}^2 + \sigma_{e0}^2))^{1/2}}$$

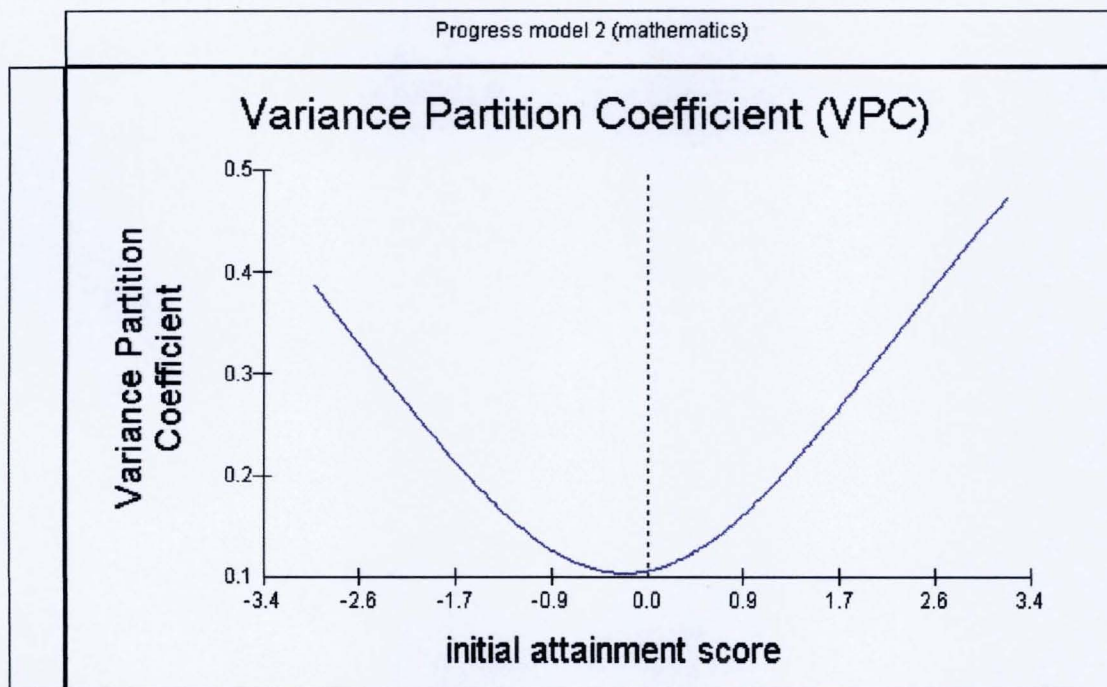
(Goldstein, 2002, p. 5)

For *mathematics*, at the average initial attainment score, the VPC becomes:

$$VPC = \frac{0.044}{0.044 + 0.282} = \frac{0.044}{0.326} = 13.5 \%$$

According to progress model 2, at the average initial attainment score, 13.5 % of the total variation in pupils' progress in mathematics can be attributed to factors *at the classroom level*.

The graph below shows that the Variance Partition Coefficient varies according to pupils' initial attainment scores. Similarly to the variance function, the value of VPC reaches a minimum of 13.23 for initial score values of approximately -0.02 and then increases for initial score values greater than -0.02 or less than -0.02 .



Graph 4: Variance Partition Coefficient by initial attainment score for mathematics, according to progress model 2.

For example, for an initial attainment one normal score unit above average, the VPC takes the value 20.7 % and for two normal score units above average the VPC takes the value 33.9 %. For values of initial score of approximately zero the VPC is practically zero. The graph shows that there is a strong classroom effect for both high-attaining and low-attaining pupils at the beginning of the year, but the classroom effect is small for pupils who initially attained at an average level.

3. 14. Evaluating the relative standing of classrooms according to a progress model with random coefficients (progress model 4) for language

For *language*, progress models with random terms associated with initial attainment score at level 2 can further be improved by allowing random terms associated with pupils' initial attainment scores at level 1, in other words by additionally allowing the variance at the pupil level (level 1) to be a quadratic function of pupils' initial attainment scores. In the derived model (*progress model 4*), pupils' initial attainment scores are allowed to vary both at the classroom level and at the pupil level.

This is known as the '*full random model*' (Goldstein (1995)):

$$y_{ij} = \beta_0 + \beta_{1j}x_{ij} + (u_{0j} + e_{0ij} + e_{1ij}x_{ij}),$$

The intercept for a classroom j β_{0j} is derived from adding the average intercept

β_0 with each classroom's intercept residual u_{0j} , which takes a different value for each classroom:

$$\beta_{0j} = \beta_0 + u_{0j}$$

and similarly the slope for a classroom j β_{1j} of x_{ij} is derived from adding the average

slope β_1 to each classroom's slope residual u_{1j} :

$$\beta_{1j} = \beta_1 + u_{1j}$$

and where

$$\text{var}(e_{0ij}) = \sigma_{e0}^2, \text{var}(e_{1ij}) = \sigma_{e1}^2, \text{cov}(e_{0ij}, e_{1ij}) = \sigma_{e01}$$

In this model the variance at both the classroom level and pupil level is modelled as a quadratic function of pupils' initial attainment scores.

The variance at the classroom level is:

$$\text{var}(u_{0j} + u_{1j}x_{1ij}) = \sigma_{u0}^2 + 2\sigma_{u01}x_{1ij} + \sigma_{u1}^2x_{1ij}^2$$

While the variance at the pupil level is:

$$\text{var}(e_{0j} + e_{1j}x_{1ij}) = \sigma_{e0}^2 + 2\sigma_{e01}x_{1ij} + \sigma_{e1}^2x_{1ij}^2$$

These variance functions stem from the fact that initial attainment score has random terms both at the classroom level as well as at the pupil level. The existence of random terms associated with pupils' initial attainment scores at level 2 means that differential progress rates are associated with different classrooms according to pupils' initial attainment score. Graph 9 shows that some of the predicted classroom lines have steeper slopes while other classroom lines have less steep slopes. Thus in some classrooms, initially high achieving pupils progress more than initially low achieving pupils, while in other classrooms all pupils progress at a fairly similar rate, no matter how well they performed initially.

The complex variation at level 1 implies that pupils' final attainment scores have non-constant variance. They are heteroscedastic and the level 1 variance is a quadratic function of initial attainment score. Pupils' final attainment scores are more variable if pupils initially performed at a high level than if they performed at a low level. In other words, for initially highly achieving pupils there is more uncertainty about their final performance, than for low achievers.

Figure 8: Progress model 4 for language

Equations

$$\text{norm_end_total}_{ij} \sim N(XB, \Omega)$$

$$\text{norm_end_total}_{ij} = \beta_{0ij}\text{constant} + \beta_{1ij}\text{norm_beg_total}_{ij}$$

$$\beta_{0ij} = -0.006(0.020) + u_{0j} + e_{0ij}$$

$$\beta_{1ij} = 0.919(0.015) + u_{1j} + e_{1ij}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} 0.023(0.005) & \\ 0.009(0.003) & 0.006(0.003) \end{bmatrix}$$

$$\begin{bmatrix} e_{0ij} \\ e_{1ij} \end{bmatrix} \sim N(0, \Omega_e) : \Omega_e = \begin{bmatrix} 0.142(0.007) & \\ 0.024(0.004) & 0.025(0.006) \end{bmatrix}$$

$$-2*\text{loglikelihood}(IGLS) = 1629.348(1503 \text{ of } 1858 \text{ cases in use})$$

Fonts Subs Name + - Add Term Estimates Nonlinear ? Help Clear

Progress model 4 (shown in figure 8 above) is an improvement on progress model 3

(2c) at $p < 0.001$, ($X^2_2 = 46.6$).

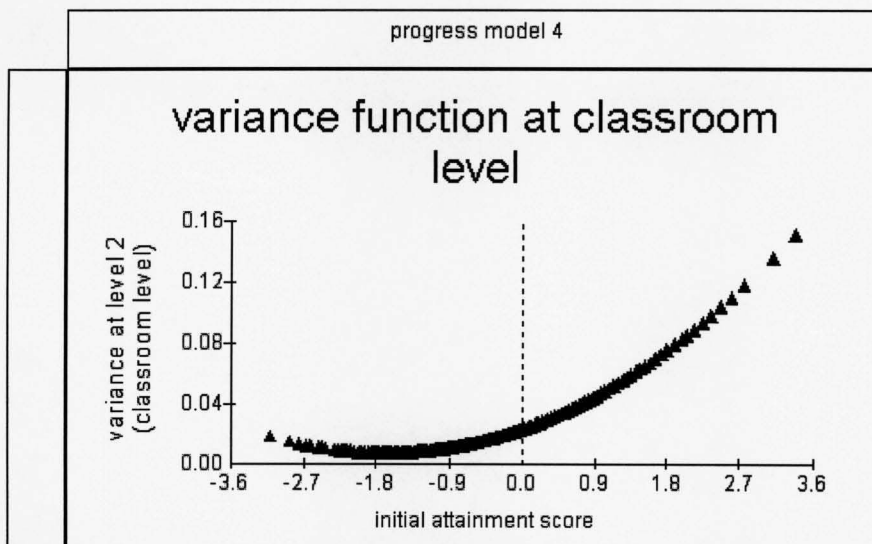
The coefficient of initial attainment score is 0.919. One normal score unit difference in pupils' initial attainment scores is associated with 0.919 normal score unit difference in final attainment scores. The coefficients of the fixed part of the model (intercept = -0.006 and slope = 0.919) represent the intercept and the slope for the average classroom. This model is the final progress model for language, upon which progress models adjusted for individual pupils' characteristics and contextual models will be based. A quadratic term involving the initial score did not fit ($p = 0.078$, $X^2_1 = 3.1$, model 2e).

3. 15. Variance function for language at the classroom level

The level-2 variance (at the classroom level) changes according to pupils' initial attainment scores, according to the following quadratic function:

$$\text{var}(u_{0j} + u_{1j}x_{1ij}) = \sigma_{u0}^2 + 2\sigma_{u01}x_0x_{1ij} + \sigma_{u1}^2x_{1ij}^2$$

The diagram below shows that classroom-level variance is an increasing function of initial attainment score. The variance function is almost zero for low values of initial attainment score (less than -0.9). Then it increases smoothly to reach a maximum value of 0.158 for an initial attainment score 3.5 normal score units above average. Thus classrooms vary very little for low achieving pupils. They vary the most for initially highly achieving pupils.



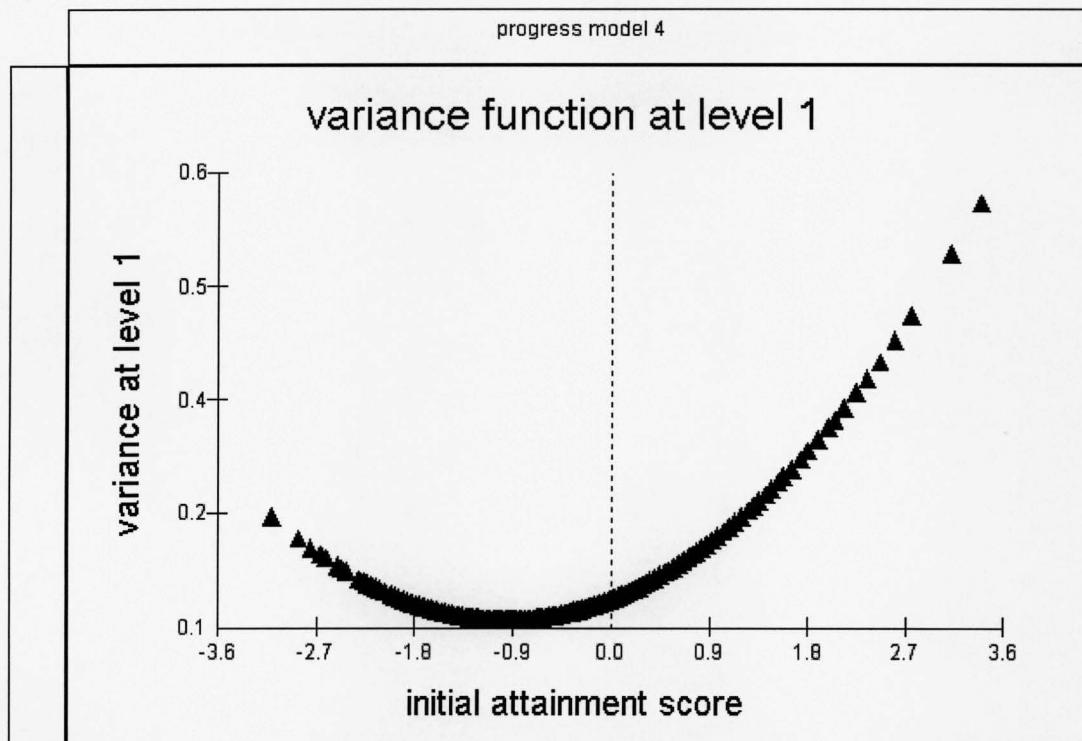
Graph 5: Level-2 variance function by initial attainment score for language according to progress model 4.

3. 16. *Variance function for language at the pupil level*

The level 1 variance is also a quadratic function of pupils' initial attainment scores according to the following equation:

$$\text{var}(e_{0j}x_0 + e_{1j}x_{1ij}) = \sigma_{e0}^2 x_0^2 + 2\sigma_{e01}x_0x_{1ij} + \sigma_{e1}^2 x_{1ij}^2$$

This variance function reaches a minimum for an initial attainment of -0.9 . Then it increases smoothly to reach a maximum value of 0.6 for values of initial attainment score 3.6 normal score units above average. At the opposite end of the attainment distribution (for initially low achieving pupils) the variance function increases very little. Hence, variation between pupils' final attainment scores is very small for pupils with an average or with a below-average initial attainment score; that is to say for values of initial attainment score less than 0.9 . There is more uncertainty in final attainment scores of initially high achieving pupils than there is for pupils who initially performed at an average or at a below-average level.



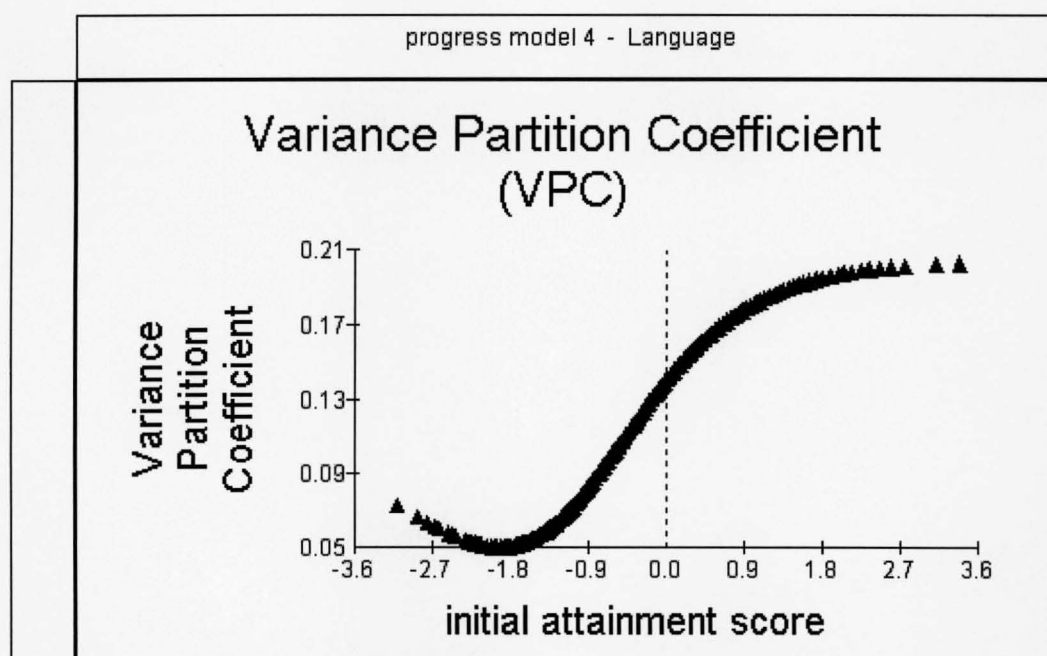
Graph 6: *Variance function at the pupil level (level 1) by initial attainment score for language according to progress model 4.*

3. 17. Variance Partition Coefficient for language

According to progress model 4, at the average initial attainment score, the Variance Partition Coefficient takes the value:

$$VPC = 13.9\%$$

At the average initial attainment score, 13.9 % of the total variation in pupils' progress in language can be attributed to classroom level factors. The graph below shows the relationship between the VPC and pupils' initial attainment scores. It shows that the VPC increases in a curvilinear fashion; it increases more steeply for initial score values from -1.8 to zero, and then it increases less steeply up to values of initial score of 1.8, where it reaches a maximum.



Graph 7: Variance Partition Coefficient by initial attainment score for language according to progress model 4.

Although the VPC increases as pupils' initial attainment scores increase, from a certain value of initial attainment score (1.8), it increases only very slightly. The VPC approximates zero for initially low-attaining pupils.

The values of VPC(s) calculated are comparable to those reported by British and American studies. To recapitulate, the current study found that 13.5% and 13.9 % of the total variation could be attributed to the classroom level at average initial score according to progress models for mathematics and language respectively.

In England, Strand (1998, p. 128) reported that at the end of Key Stage 2 the percentage of total variance that could be attributed to classroom membership for progress models was 11.4% for mathematics, and 14.5% for language.

Taking into account that in the current study the period between the two measurements was six months, it may be concluded that for Greek state primary schools in Piraeus, classrooms matter but not as much as schools matter in Britain. Furthermore classrooms have a similar impact in terms of pupils' progress in both subjects (language and mathematics). However, for language, classrooms add more to the progress of initially high achieving pupils. Also although for progress models 2 for mathematics and progress model 4 for language the values of VPC for mathematics and language are close, the values of VPC derived from null models are further apart, as has been previously shown.

3. 18. Predictions according to progress model 2 for mathematics

Below, final attainment scores for different values of pupils' initial attainment scores will be predicted for different classrooms. Both initial attainment and final attainment scores are expressed in normal score units. The predictions will be based on *Progress model 2*.

Progress model 2 has been derived from the following equations:

$$y_{ij} = \beta_{0j} + \beta_{1j}x_{ij}$$

In which

$$\beta_{0j} = \beta_0 + u_{0j}$$

is the intercept for each classroom;

and

$$\beta_{1j} = \beta_1 + u_{1j}$$

is the slope for each classroom.

In this case y_{ij} is pupils' final attainment score, which is a function of pupils' initial attainment score x_{ij} , β_0 is the average intercept, and u_{0j} is the residual term for each classroom. The latter term plays the role of a differential intercept for each classroom.

β_1 is the average slope across classrooms, and u_{1j} is a classroom residual signifying the departure between a classroom slope from the average classroom slope.

If the values of intercept and slope coefficients derived from progress model 2 are introduced in the above model, we get:

$$\hat{y}_{ij} = 0.048 + (0.802 + \hat{u}_{1j})x_{ij} + \hat{u}_{0j}$$

According to this model, the final attainment score for each classroom depends linearly on pupils' initial attainment scores, but the predicted final attainment score for each classroom is derived through a different within class regression. In each classroom's regression equation the intercepts are estimated by adding each classroom's intercept residual u_{0j} to the average intercept (0.048) while each classroom's slope is estimated by adding the average slope (0.802) to the classroom's slope residual u_{1j} .

a) Example 1 in mathematics:

For one school in Piraeus, the *first classroom* was found to perform at an above-average level while the second classroom was found to perform at an average level.

The intercept and slope residuals for the first classroom are:

For the first classroom

Intercept: -0.044, Slope: 0.047

While for the second classroom:

Intercept: -0.155, Slope: 0.073

If these intercept and slope residuals are introduced in the above equation, the following within-classroom regression is obtained for the first classroom:

$$\hat{y}_{i1} = 0.004 + (0.849)x_{i1}$$

And the following within classroom regression is obtained for the second classroom:

$$\hat{y}_{i2} = -0.107 + (0.875)x_{i2}$$

If different values of initial score are introduced in the above equation, the following predicted final attainment scores are obtained:

For the *first classroom*, the predicted final attainment score for pupils who performed two normal score units above average, is:

$$\hat{y}_{i1} = 0.004 + 0.849 \cdot 2 = 1.702$$

For the *second classroom*, the predicted end score for pupils who performed two normal score units above average is:

$$\hat{y}_{i2} = -0.107 + 0.875 \cdot 2 = 1.643$$

In the following table (3. 9.) all the relevant calculations are summarised in order to facilitate classroom comparisons based on pupils' predicted final attainment scores.

Table 3. 9. Predictions for low, average and high achievers according to progress model 2 for two classrooms belonging to the same school

Pupils' Initial Attainment Score:	Predicted end scores for:		
	1 st classroom (Performing at an above-average level).	2 nd classroom belonging to the same school (This classroom performs at an average level).	Difference in predicted end scores between the 2 classrooms
2 normal score units above average	1.702	1.643	0.059
1 normal score unit above average	0.862	0.768	0.094
Average initial score	0.004	-0.107	0.111
1 normal score unit below average	-0.845	-0.982	0.137
2 normal score units below average	-1.694	-1.857	0.163

The two classrooms in the examined school differ more for initially low attaining pupils than for initially average or high attaining pupils.

b) Example 2 in mathematics:

For another school with two classrooms, the first classroom was found to perform at a below-average level whereas the second classroom was found to perform at an average level.

The intercept and slope residuals for the first classroom are:

Intercept: -0.335, Slope: -0.057

Whereas, for the second classroom:

Intercept: -0.073, Slope: 0.129

In a similar way, one can predict the final attainment scores for initially high achieving, average achieving and low achieving pupils.

Thus the following table summarises the predictions for the two classrooms belonging to this second school:

Table 3. 10: Predictions for low achieving, average achieving and high achieving pupils for two classrooms belonging to the same school. All predictions are based on progress model 2.

	Predicted final attainment scores for:		
Pupils' Initial Attainment scores:	1st classroom (This classroom performs at a below-average level)	2nd classroom belonging to the same school. (This classroom performs at an average level)	Difference in pupils' predicted final attainment scores between the two classrooms
2 normal score units above average	1.203	1.837	-0.634
1 normal score unit above average	0.458	0.906	-0.448
Average initial score	-0.287	-0.025	-0.262
1 normal score unit below average	-1.032	-0.956	-0.076
2 normal score units below average	-1.777	-1.887	0.11

In the second school's case it is evident that the difference in predicted final attainment scores between the two classrooms is bigger for initially high achieving pupils than for initially average or low achieving pupils.

If pupils' predicted final attainment scores derived from these two schools are compared, then it can be noticed that there is no consistent pattern in the way classrooms within the same school differ. Whereas for the 1st school the biggest difference in predicted final attainment scores between the two classrooms was for low achieving pupils, for the 2nd school, the biggest difference was for high achieving pupils.

3. 19. Predictions according to progress model 4 for language

In the case of language, pupils' predicted final attainment scores for different values of initial attainment score will be calculated using 'progress model 4', since this is the final progress model fitted. Both initial attainment and final attainment scores are expressed in normal score units. In progress model 4 pupils' initial attainment is included as the only explanatory variable, random at both classroom level and pupil level.

Progress model 4 has been derived with the following equations:

$$y_{ij} = (\beta_0 + u_{0j}) + (\beta_{1j} + u_{1j})x_{1ij} + e_{ij}$$

In which:

$$\beta_{0j} = \beta_0 + u_{0j}$$

$$\beta_{1j} = \beta_1 + u_{1j}$$

is the intercept for classroom j;

is the slope of x_{1ij} in classroom j.

The values of the intercept and slope coefficients estimated from progress model 4 lead to the following regression equation:

$$\hat{y}_{ij} = -0.006 + (0.919 + \hat{u}_{1j})x_{ij} + \hat{u}_{0j}$$

According to this equation, pupils' final attainment score for each classroom depends linearly on pupils' initial attainment score, but the predicted final attainment score is

derived through a different within-class regression. In order that the within-class regressions for each classroom be estimated, the intercept for each classroom is formed by adding -0.006 to the classroom level intercept residual u_{0j} , whereas, the slope for each classroom is formed by adding 0.919 to the classroom's slope residual u_{1j} .

One example for language:

For another school in Piraeus in which two 6th grade classrooms operate, *the first classroom was found to perform at an average level while the second classroom was found to perform at an above-average level in terms of progress*. The difference between the rankings of their intercept residuals was 67 positions. The intercept and slope residuals for the first classroom were:

Intercept: -0.141, Slope: -0.088

And for the second classroom they were:

Intercept: 0.204, Slope: 0.101

One can predict pupils' final attainment score for initially low achieving, average achieving and high achieving pupils enrolled in these two classrooms.

The following table summarises the final attainment score predictions for pupils enrolled in these two classrooms:

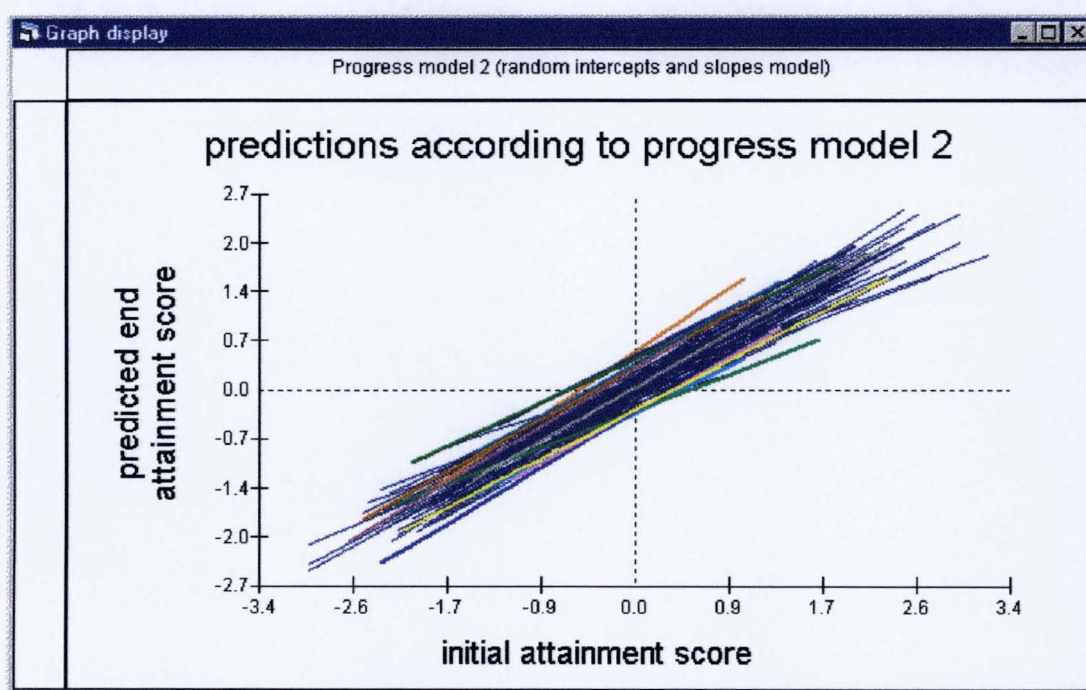
Table 3. 11: Predictions for low, average and high achievers according to progress model 4 for two classrooms belonging to the same school

Pupils' Initial Attainment scores:	Predicted final attainment scores for:		
	1st classroom (Performing at an above-average level).	2nd classroom in the same school. (This classroom performs at an average level).	Difference in predicted final attainment scores between the two classrooms
2 normal score units above average	2.24	1.515	0.725
1 normal score unit above average	1.219	0.684	0.535
Average initial attainment score	0.198	-0.147	0.345
1 normal score unit below average	-0.82	-0.98	0.16
2 normal score units below average	-1.84	-1.81	-0.03

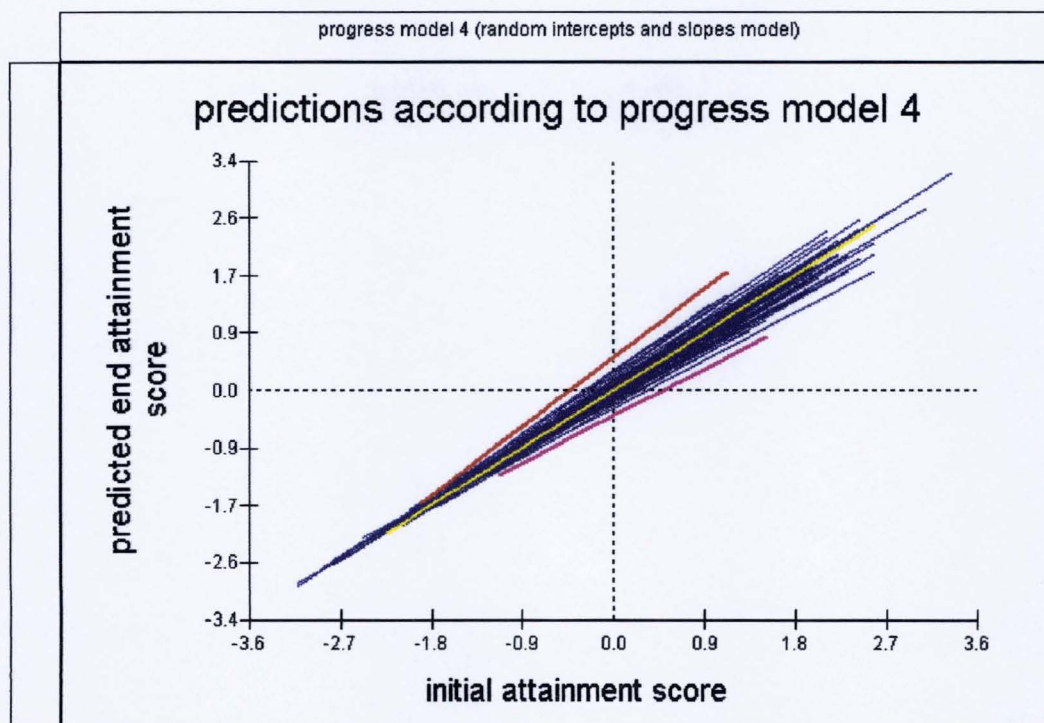
When examining the difference between pupils' predicted final attainment scores it is obvious that the first classroom generally performs at a higher level, with the exception of pupils who initially performed well below average. The biggest difference in predicted final attainment scores between these two classrooms is for pupils who initially attained well above average. However, the advantage associated with being enrolled in the first classroom diminishes as pupils' initial attainment scores decrease. Finally pupils with well below average initial attainment score are better off in the second classroom. In that case only the second classroom promises a brighter future to these very low achieving pupils. This finding can be demonstrated for other schools as well, if the differences in pupils' final attainment scores between two classrooms belonging to the same school are calculated.

**3. 20. Predictions according to progress model 2 for mathematics and
progress model 4 for language**

In graphs 8 and 9 below classroom lines relating pupils' predicted final attainment scores with their initial attainment scores are shown. These graphs are derived from 'progress model 2' for mathematics, and from 'progress model 4' for language.



Graph 8: Final attainment score predictions for each classroom according to 'progress' model 2 for mathematics.



Graph 9: Final attainment score predictions according to 'progress model 4' for language

Significant classroom variation in relation to pupils' initial attainment scores is identified for both language and mathematics. Both the above graphs show that classroom lines vary in both their intercepts and in their slopes. 'Progress model 2' for mathematics and 'progress model 4' for language allow for random terms associated with pupils' initial attainment scores at the classroom level. The fact that classroom lines cross over means that classrooms contribute differently to the progress made for pupils initially performing at an above-average, average or below-average level. Differential effects in relation to pupils' initial attainment scores exist. Graphs 8 and 9 demonstrate the relationship between initial attainment score and final attainment score, which varies according to the classroom attended.

In mathematics, to some extent, some classrooms boost the progress of initially high achieving pupils more than the progress of initially low achieving pupils, whereas other

classrooms promote more equally the performance of low achieving, average achieving and high achieving pupils. However, for language differential classroom effects are more apparent for initially high achieving pupils than average achievers, whereas hardly any classroom effects are identified for low achieving pupils. For language there is less variation between classrooms for average achieving pupils than for high achieving pupils and practically no variation for low achieving pupils. This is shown by the fact that classroom lines are very close to one another for low values of initial score and gradually fan out for average to high values of initial score. This aspect was also highlighted in graph 6 that shows the variance function at the classroom level for language according to progress model 4. This variance function was shown to be almost zero for values of initial score less than 0.9 normal score units and then it gradually increased, up to around 0.16.

Suppose that the slope (final attainment score / beginning attainment score) was stable for all classrooms or that the slope residuals did not exist. Then all classrooms would have contributed the same amount of final attainment score for an extra unit of initial attainment score. In the prediction graph, classroom lines showing the relationship between pupils' initial attainment scores and their final attainment scores would have been parallel (as in graph 1) instead of crossing over. The fact that classroom lines cross over shows the existence of differential effects.

3. 21. *Investigating the 'school effect'*

In further stages of modelling and for both subjects, null models were created in which the overall variance was divided into three levels (between schools, classrooms and pupils). These 'null models 3' did not control for any additional variables and turned out not to be an improvement on null models 2 that partitioned the total variation into the classroom level and the pupil level: $X_1^2 = 0.9$, $p = 0.342$ for mathematics and

$X_1^2 = 1.3$, $p = 0.254$ for language.

Luyten (1994, p. 87) suggested that it is possible for school effects to appear after having controlled for some relevant variables. Modelling was pursued further in order to investigate whether school effects emerge after controlling for additional variables.

Scheerens and Bosker (1997) consider growth in student achievement over time as the most appropriate criterion for assessing the magnitude of school effects, since some researchers believe that school effects models are better specified using longitudinal designs. For example Goldstein (1997) stated that in order to make satisfactory inferences about individual school performances school effectiveness studies should be longitudinal so that pre-existing student differences and subsequent contingent events among institutions can be taken into account.

Therefore, the analyst examined unadjusted progress models as well as progress models adjusting for individual characteristics partitioning the total variability into three levels in order to estimate whether school effects existed. The researcher looked for a significant improvement of the model fit demonstrated in the deviance derived from the dif-

ference between the two -2loglikelihood statistics when a model with three levels is compared with its respective model with two levels. The deviance test was tested with one degree of freedom. The models partitioning the total variation into three levels (namely into pupil classroom and school levels) estimated at the same time the variance at the school level $\sigma_{\nu 0}^2$, the variance at the classroom level $\sigma_{u 0}^2$ and the variance at the pupil level $\sigma_{e 0}^2$. The analyst examined unadjusted progress models and progress models adjusting for individual characteristics in order to estimate whether school effects exist. However, none of these models turned out to be an improvement on the corresponding progress model with two levels at $p < 0.05$.

For mathematics and for language, null models 2 partition the total variation identified in pupils' final attainment scores into variation at the classroom level, and variation at the pupil level. Both Null models 2 are improvements on respective null models 1 at $p < 0.001$. For mathematics, the deviance statistic for null model 2 in comparison to null model 1 is significant. For mathematics, $X_1^2 = 75.4$, $p < 0.001$, while for language, $X_1^2 = 38.5$, $p < 0.001$).

For mathematics, classroom level variance takes the value of $\sigma_{u 0}^2 = 0.104$ (SE = 0.024, $p < 0.001$), which suggests that significant differences between classrooms in average mathematics performance exist. Pupil level variance is estimated to be $\sigma_{e 0}^2 = 0.892$, (SE = 0.032, $p < 0.001$). So VPC amounts to 0.104 (10.4%) for null model.

For mathematics, 10.4% of the total variation can be attributed to the classroom level, which implies that classroom effects exist. Classroom pooled unadjusted final attainment scores significantly vary from one classroom to the next.

The null model with three levels, which partitioned the total variation between the school level, the classroom level and the pupil level, was not an improvement on *null model 2*, which employs classroom and pupil levels only, ($X_1^2 = 0.9$, $p = 0.342$). However, even in the hypothetical situation where the null model with three levels was an improvement on null model 2 (which splits the total variability into two levels), variation at the classroom level would still account for 7.3 % of the total variability among pupils' outcomes (that would be the value for the VPC estimating the proportion of classroom-level variance to the total variance). In that hypothetical situation, the VPC for level 3 (variation at the school level) in relation to total variation would have taken the value of 3.2 % only.

For language, classroom-level variation in null model 2 is estimated to be $\sigma_{u0}^2 = 0.068$, (SE= 0.018), which suggests that significant differences between classrooms in terms of unadjusted pupils' final attainment scores exist. Variation at the pupil level is estimated to be $\sigma_{e0}^2 = 0.930$, (SE=0.034). The VPC suggests that variation at the classroom level amounts to 7% of the total variation.

For language, the null model with three levels (null model 3), which partitioned the total between the school level, the classroom level and the pupil level, was not an improvement on *null model 2*, which employs classroom and pupil levels only,

($X^2_1=1.3$, $p=0.254$). However, if null model 3 had been an improvement on null model 2, classroom level variation would account for 4.7 % of the total variation among pupils' outcomes (that would be the value for the VPC estimating the proportion of classroom-level variance to the total variance). In that hypothetical situation, the VPC for level 3 (variation at the school level) in relation to total variation would have taken the value of 2.3 % only.

The values of VPC derived from null model 2 for language (0.007) is similar to the related value of VPC for mathematics (0.073). The proximity between the values of VPC suggests that classroom effect is similar for the two subjects (for mathematics and for language).

According to the current study, in primary schools with two 6th grade classrooms, no consistent pattern emerged in relation to classroom rankings within schools. According to the unadjusted absolute attainment model (null model) the *percentage of schools having 'similar' classrooms* (away from each other's rankings, 10 positions or less) is 33% for mathematics and 29.6% for language. *For mathematics*, according to the '*adjusted attainment model controlling for pupils' ethnicity and gender*' this percentage is 40% and according to progress model 2 this percentage is 11%. According to an adjusted progress model additionally controlling for pupils' ethnic status and gender this percentage is 14.8%.

For language according to the '*adjusted absolute attainment model controlling for ethnicity and gender*' this percentage is 37% while according to '*progress model 4*' this

percentage is 26%. According to an adjusted progress model controlling for gender this percentage is 22.2%.

In addition, for both subjects and for schools with two classrooms, whilst one classroom within a given school scored at an above-average or at a below-average level (in terms of attainment, or in terms of progress, or in terms of adjusted progress) in relation to the majority of other classrooms in the city of Piraeus, the 2nd classroom belonging to the same school scored at an average level, in most cases²².

In order to make a judgement about the absence of school effects identified in the current setting, the reader should note the values of variance partition coefficients for the school level reported by international studies. Creemers et al. (2002, p. 45), in describing the international School Effectiveness Research Project reported that the percentage of the total variance that can be attributed to the school level (VPC) was for UK 0.21 and for USA 0.35, according to a progress model accounting for pupils' pre-test at entry, while the dependent variable was pupils' attainment score at the end of the first grade of primary school. School effects derived from models additionally controlling for pupils' background characteristics are 0.29 for the USA and 0.11 for UK, for the 1st grade of primary school, while school effects corresponding to progress models adjusting for individual pupils' characteristics are 0.25 for the USA and 0.10 for UK.

A recent study of departmental effectiveness in secondary schools in UK (Thomas, Sammons & Mortimore, 1995) found that the percentage of total variance attributed to the school level in terms of progress was lower for English (6%) and higher for mathe-

22 This information is shown in the 3rd appendix.

matics (9%). Sammons and Mortimore (*op. cit.*) explain this on the premise that English was also a subject influenced by the pupils' experiences outside the school while mathematics is a subject primarily taught in school.

To summarise, although null models as well as absolute attainment and progress models with three levels were considered in order to examine the possibility that school effects might exist in the current dataset, none of the models considered was a significant improvement on the corresponding model with two levels. Hence no school effects were identified among the 6th grades of this Greek state primary school setting. The absence of school effects is discussed in the findings, in section 4. 2.

3. 22. Comparing classroom raw and normalised final performance levels

Simple arithmetic means for classroom total mean raw end scores, varied between 33.1 and 68.7 for mathematics and 76.5 and 133.3 for language respectively. Pupils' scores could range from 1 to 93 for mathematics and from 1 to 160 for language.

Table 3. 12: Range in classroom mean raw scores at the end of the 6th grade.

	Pupils' Mean	Standard deviation (At the pupil level)	Lowest Mean Raw Score	Highest Mean Row Score
Mathematics	46.9	26	28.6	64.3
Language	105.5	33.4	76.5	133.4

In mathematics pupils attained an overall average score of 46.9 with a standard deviation of 26. In language pupils attained an overall average score of 105.5 with a standard deviation of 33.4. The two classrooms with highest and lowest pupils' final attainment scores differed by 35.7 raw score units for mathematics and by 56.9 raw score units for language. For mathematics this difference in magnitude is approximately half the score attained by the best scoring classroom.

However, raw results should not be employed in modelling, as they are not comparable between subjects, as the scales used to measure pupils' initial and final attainment scores for mathematics and for language have different lengths. One unit increase in pupils' initial attainment scores (the main explanatory variable) would not be equal to one unit increase in the dependent variable. In addition, if raw results were used, fixed part coefficients or random part parameters would not be comparable across subjects.

To rectify this, *normalised test scores*²³ have been employed instead of raw attainment scores so that all measures are expressed on the same scale. Normalised test scores have zero means and a unit standard deviation. This transformation aims to put all measures on the same scale.

Table 3. 13: Range in classroom mean normalized scores at the end of the 6th grade.

	Pupils' Mean	Standard deviation (At the pupil level)	Lowest Mean Row Score	Highest Mean Raw Score
Mathematics	0	1	0.6	3.5
Language	0	1	-0.8	1.1

For mathematics, in terms of classroom average normalised end scores the means of the lowest and the highest scoring classrooms were 0.6 and 3.5 (where the obtained range at the pupil level was from -3.2 to 3.1 with a mean score around zero and standard deviation 0.997, which by definition is almost one normal score unit). For language, in terms of classroom average normalised final attainment scores, the means of the lowest and the highest scoring classrooms were -0.8 and 1.1. The classrooms with the highest and the lowest mean final attainment scores differed by 2.9 normal-score units for mathematics and by 1.9 normal-score units for language.

²³ *Normal Scores transformation*: A method according to which expected values from the standard normal distribution are assigned instead of the original scores, according to the ranks of the original scores. The output column contains the Normal Equivalent Deviates (NED) of $(i-0.5/n)$ where i ranks the values in the input column and n is the number of values (Rashbash and Woodhouse, 1996 in Kaluge, 1998, p. 124).

3. 23. Classroom residuals

The difference in terms of classroom residual estimates derived from null model ranges from -0.6 to 0.6. There is great discrepancy between the average classroom normal scores and the classroom residuals expressed in normal score units derived from a null multilevel model. Classroom residuals calculated through multilevel modelling techniques are affected by a ‘shrinkage factor’. The shrunken residuals are calculated as:

$$\hat{u}_j = \left(\frac{n_j \sigma_u^2}{n_j \sigma_u^2 + \sigma_e^2} \right) \hat{y}_j, \quad \hat{y} = \frac{1}{n_j} \sum_i (y_{ij} - \beta_0 - \beta_1 x_{ij})$$

(Goldstein, 1997, p. 380)

Where

\hat{y}_j is the classroom residual and

Where the shrinkage factor is the term in brackets multiplying \hat{y}_{ij} .

According to this formula the classroom residual \hat{y} (derived from the simple average of the deviations of the fixed part prediction of the model $\beta_0 + \beta_1 x_{ij}$ from pupils’ observed scores in the classroom j y_{ij} derived from the second equation), is multiplied by a shrinkage factor (within brackets in the first equation). As the number of pupils in classrooms for which classroom residuals are estimated increases, the shrinkage factor tends

to become one. If the shrinkage factor approaches one, it affects only very little the unadjusted classroom residual \tilde{y} . On the other hand, if the class size (number of pupils in a given classroom j) is small, the shrinkage factor pulls the classroom residual towards the average. *‘Increased shrinkage for a small level 2 unit can be regarded as expressing the relative lack of information in the unit so that the best estimate places the predicted residual close to the overall population value as given by the fixed part’* (Goldstein, 1995, p. 24). Thus whereas this ‘shrinkage’ factor pulls the residuals of classrooms with small size towards the average it tends not to affect the residuals of big classrooms. Ultimately the shrinkage factor takes the value of one as the number of pupils within a given classroom increases. *‘Increased shrinkage for a small level 2 unit can be regarded as expressing the relative lack of information in the unit so that the best estimate places the predicted residual close to the overall population value as given by the fixed part’* (Goldstein, 1995, p. 24). The ‘shrinkage’ factor pulls small classroom raw residuals towards the average but tends not to affect the value of the raw residuals of big classrooms. *‘The accuracy of these estimates will depend largely on the number of students in each school’* (Goldstein, 1997). In addition, the shrinkage factor is inversely related to variance at the pupil level (level 1) and is inversely proportionally related to pupil level variance. As the latter decreases, the shrinkage factor increases and approaches one. For large values of pupil-level variance the shrinkage factor decreases.

3. 24. Classroom effects

Many School Effectiveness studies focus on the school effect or the classroom effect, which can be described as *the relative impact of the school or of the classroom* in improving pupils' learning outcomes. Classroom effects point to differences between classrooms within schools, while school effects point to differences between schools.

Since school effects were not identified, the VPC (Goldstein, 2002, p. 3) shows the relative importance of the level 2, which in the current study is the classroom level variance, compared with the total variance. The total variance is the sum of variances at the classroom level and at the pupil level.

Rowe et al. (1995, cited in Goldstein, 1997, p. 390) point out that since learning takes place in classrooms as well as in schools, the classroom level is a key one to model. Hill and Rowe (1996, cited in Scheerens and Bosker, 1997) suggested that teacher effectiveness is a more probable cause of differences between schools than the schools' effectiveness itself. *'The classroom level has maybe two or three times more influence on student achievement than the school level does'* (Creemers, 1994). *'Shifting the emphasis from the school to the classroom level is based on the thought that learning takes place in classrooms, and that factors at higher levels are of a conditional kind'* (Creemers et al., 1997).

Luyten (2003, p. 31), in his overview of the research literature about the size of school effects compared to teacher (or classroom) effects, shows that *'there appears to be a general consensus that teacher effects outweigh school effects'*. Scheerens and Bosker (1997, p. 96) conclude that teacher effects predominate over school effects and Teddlie

and Reynolds (2000, p. 158) take a similar view. Hill and Rowe (1996) concluded that schools do make a difference, but it is mainly teachers that cause schools to differ, while the unique effect that schools add to teacher effect seems relatively small. In his meta-analysis of 16 School Effectiveness studies Luyten (*op. cit.*) reports that larger teacher effects have been found in primary than in secondary schools for both language and mathematics (p. 45).

Although according to Reynolds et al. (1996) a substantial part of SER has been characterised as sociologically inspired, *'some critics have called for SER to focus more closely on classroom processes, accompanying our understanding of differential achievement with a systematic analysis of learning and teaching'* (Robertson and Toal., cited in Macbeath and Mortimore, 2001, p. 107).

'More important, what interests teachers the most is what goes on between them and their students. Placing a central focus on this is therefore of fundamental importance for successful schools. This means focusing on teaching and, particularly, learning – the treasure within' (Delors et al., 1996, cited in Stoll, 1999, p. 504).

There are a number of models that can be drawn for the conduct of process-product research and the reader is referred to Dunkin & Biddle (1974), Bennett (1996), Creemers (1994) and Willms (1992) (authors cited in Goldstein and Blatchford, 1998). However, among the aims of the current study is not to highlight conditions that boost pupils' outcomes at the classroom level, as a process-product study would do.

The fact that that classroom effects outweigh school effects has important implications regarding the way in which the results from School Effectiveness analysis should be reported. It should be made clear whether the study looked for the existence of classroom effects in null, absolute-attainment and progress models. Such models partitioning the variation into three levels have to be created and tested, and the variances corresponding to each of the three levels should be reported. Rowe and Hill (1994, cited in Scheerens and Bosker, 1997, p. 80) state that including the classroom level in a model already including the school level and the pupil level would lead to a decline in school effect size. The authors attribute this decline to the fact that the model does not include all levels, which leads to an artificial statistical increase in the size of school effects.

In the current study fairly similar results in terms of the percentage of total variance attributed to the classroom level are derived from progress models 2 and 4 for mathematics (13.5) and language (13.9) respectively. It seems that in the Greek primary setting classrooms influence pupils' progress similarly for both subjects. Classroom effects were smaller for null models. The intra class correlation coefficients were estimated to be 10.4% for mathematics and 6.8% for language according to null models 2.

In school effectiveness literature the individual values of the residuals u_{0j} are usually interpreted as the 'effect' associated with the level-2 unit, which in many British and American School Effectiveness studies represents the school level. In both the British and American literature the existence of significant variance at level 2 indicates *school effects*. The current study tested whether school and classroom effects exist by creating models partitioning the total variance between school, classroom and pupil levels.

However, as models including three levels were not significant improvements on respective models with two levels in none of the models examined, it can be deduced that for state schools in the city of Piraeus only classroom effects exist. For this reason, the level-3 (school-level) variance was removed from all subsequent phases of modelling. Thus absolute attainment models, progress models and contextual models included only two levels (pupil level and classroom level).

Opdenakker and Van Damme (2000) explored the effects of ignoring top or intermediate levels on the fixed effects and variances in the random intercept models. Their results show that *'ignoring one or more top levels causes an overestimation of the variance belonging to the highest level considered. The variance of the other levels is unaffected'* (p. 121). In the present analysis when school-level variance is removed from the models, classroom-level variance terms may be slightly but not substantially inflated. When the variance term at the school level (level 3) was removed, a small fraction of the variation previously allocated at the school level was reassigned to classroom (2nd) level. In the current study, in order to investigate how the elimination of the school-level influences the classroom-level (level 2) variance, progress models with and without the school-level variance term were compared. It was found that variation at the classroom level increases by 0.010 units (from 0.030 to 0.040 SE=0.009) when variation at the school level is removed from the progress model. Nevertheless, before removing the school-level variance, classroom-level variance was greater than twice its standard error. For mathematics, the progress model in which the total variation is divided into three levels is not an improvement on the progress model with two levels ($p=0.4$). Thus, removing the school-level variance does not artificially produce a class-

room effect that has not already been there. Classroom effects are already present in the 3-level model and removing the school-level variance term from progress models leads to a slight increase in the level-2 variance term. However variance at level-2 (classrooms) was significant in all absolute attainment and progress models with three levels.

3. 25. Examining the classroom residuals derived from null and progress models

Graphs 10, 11, and the upper part of graphs 12 and 13 show classroom residuals. These residuals represent classroom relative rankings according to a given criterion, which in the case of the null models is pupils' unadjusted final attainment scores, while for progress models it is pupils' progress during a pre-specified period of time.

The vertical lines are 95% confidence intervals. Confidence intervals associated with these residuals show whether a given classroom differs in terms of the criterion of interest from the average. *'There is the need to interpret residual estimates of an individual school's effects by reference to the confidence limits associated with such estimates'* (Creemers, 1994a; Goldstein et al., 1993). The fewer the pupils in a given classroom, the larger the confidence interval for the classroom's residual. Classrooms with overlapping confidence intervals have scored similarly in terms of final attainment or in terms of progress. From graphs 10 to 13 it is evident that most of the classrooms have similar predicted final attainment scores and that most of the classrooms have made similar progress. Since the confidence intervals attached to the intercept and slope residuals of the majority of the classrooms investigated touch the average line, comparisons between individual classrooms are not legitimate; instead classrooms can be compared with the average. It is clearly understood that *'even when suitable adjustments for*

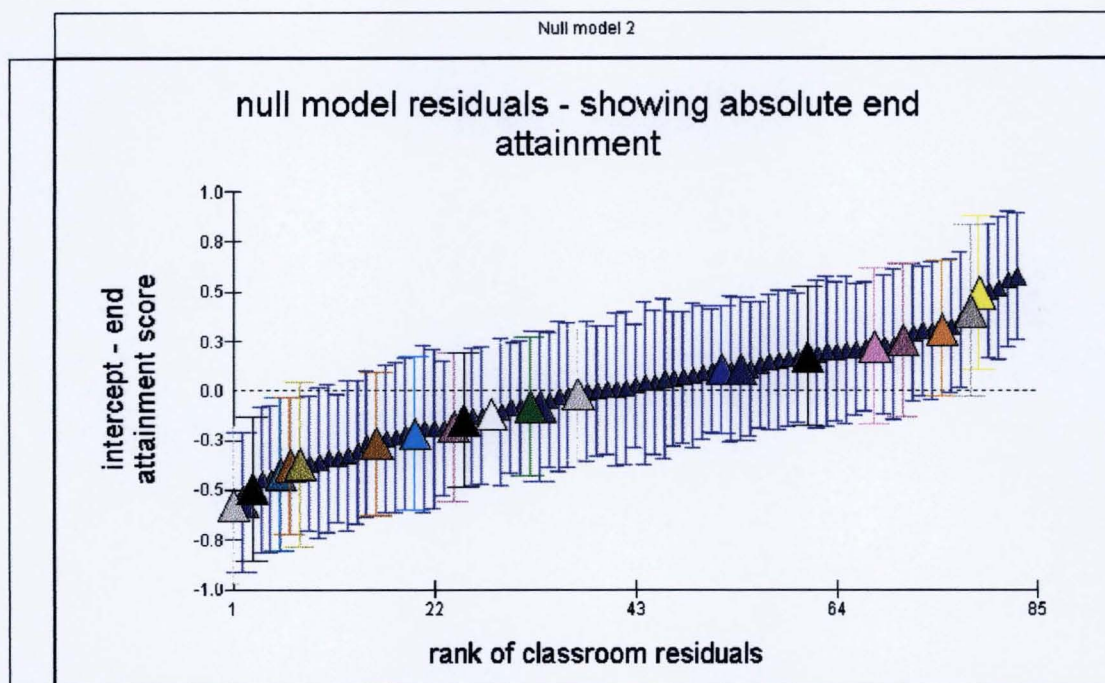
intake achievement and other relevant factors are done, the resulting value added estimates usually have too much uncertainty attached to them to provide reliable rankings' (Goldstein 1997).

However, graphs 10 and 11 show that there are a few classrooms with final attainment scores significantly above or below average and graphs 12 and 13 show that there are few classrooms with progress rates significantly above average or below average.

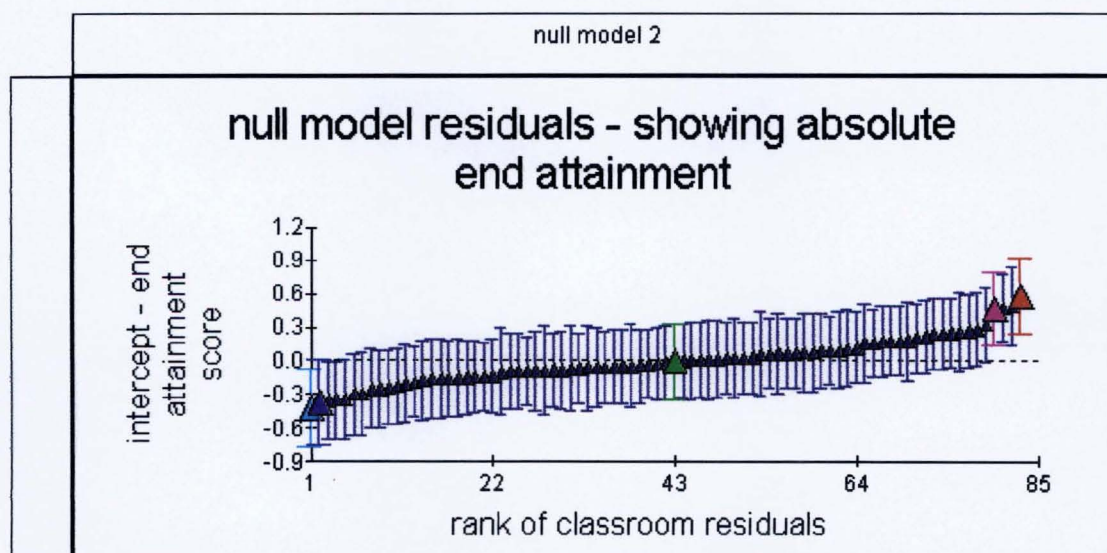
Classrooms with confidence intervals crossing the horizontal line are considered to have scored at an average level according to the criterion of interest (in terms of final attainment score or in terms of progress). Classrooms with confidence intervals above the horizontal line scored significantly better than average, whereas those with confidence intervals below the horizontal line scored significantly worse than average.

According to Goldstein & Spiegelhalten (1996), by taking account of sampling variability associated with residual estimates it is possible to establish whether the level 2 units (classrooms) scored significantly better or worse (at the 0.05 level) than others.

3. 26. Comparing classroom residuals derived from null models



Graph 10: Classrooms ranked according to null model 2 in mathematics (including classroom and pupil levels). 95% confidence intervals are applied.



Graph 11: Classrooms ranked according to null model 2 for language (including classroom and pupil levels). 95% confidence intervals are applied.

The above graphs show classroom residuals of the respective null models 2 for mathematics and language ranked by magnitude.

According to Sheerens and Bosker (1997, p. 70) residuals derived from a null model, which does not take account of pupils' characteristics represent '*gross classroom effects*'. Graphs 10 and 11 show classroom relative rankings according to pupils' unadjusted final attainment scores for mathematics and language.

Graph 10 shows that from the total of 83 classrooms there are 6 classrooms (7.2%) in which pupils attained an above-average level, and 9 classrooms (10.8%) in which pupils attained a below-average level. The attainment of the remaining classrooms (81.2%) did not statistically differ from the average.

The classroom that performed the highest had a confidence interval from 0.255 to 0.897 points, according to the null model. The 6th classroom, the lowest of the above average classrooms had a confidence interval from 0.016 to 0.697 points. A typically average classroom has a confidence interval that is from -0.377 to 0.392 points. The classroom that performed the lowest from those at a below-average level had a confidence interval from -0.968 to -0.182 points. The 9th classroom below average, the classroom scoring highest from those that scored at a below-average level, had a confidence interval from -0.708 to -0.028 points. The next classroom's confidence interval touches the average horizontal line. Although this classroom does not significantly differ from the average it does not differ a lot from the previous classroom, as its confidence interval is from -0.739 to 0.015 points. It is evident that there are no clear dividing lines between class-

rooms that performed at a below-average level and classrooms with adjacent rankings, which performed at an average level, given that these classroom confidence intervals slightly touch the average line. Nor are there any clear dividing lines between classrooms that performed at an above-average level and classrooms with adjacent rankings, which performed at an average level.

Graph 11 shows that for language, from a total of 83 classrooms, there are 4 classrooms (4.8%) performing at an above-average level and 4 classrooms (4.8%) performing at a below-average level in absolute terms at the end of year 6. The rest of the classrooms (90.3%) were found to perform at an average level.

3. 27. Comparing classroom residuals derived from progress models

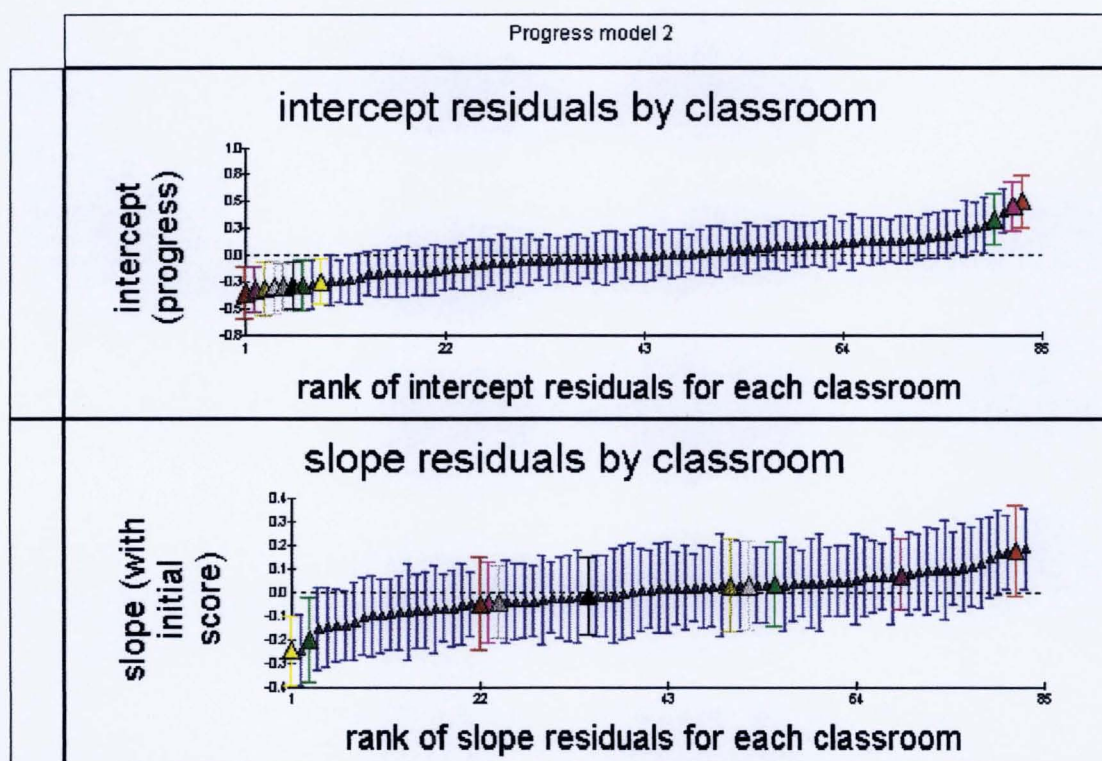
Graphs 12 and 13 show the intercept and slope residuals of progress model 2 for mathematics and progress model 4 for language (which are random coefficients models).

The upper part of graphs 12 and 13 compare classroom intercept residuals, thus enabling comparisons between classrooms in terms of progress. They demonstrate the degree to which a given classroom has progress rates different from the average. According to Sheerens and Bosker (1997, p. 70) these classroom residuals represent the impact of classrooms after controlling for pupils' initial attainment score and they represent net classroom effects.

The intercept residuals shown on the top of graphs 12 and 13 demonstrate the degree to which a given classroom has progress rates statistically different from the average, that is to say which between the classrooms investigated progressed at an above-average level, which ones progressed at an average level and which ones progressed at a below-average level.

According to Sheerens and Bosker (1997, p. 70) these classroom residuals represent the impact of classrooms after controlling for pupils' initial attainment scores and they represent net classroom effects.

The bottom part of graphs 12 and 13 demonstrate classroom slope residuals according to pupils' initial attainment scores. These slope residuals show the degree to which classrooms contribute to different progress rates for high versus low achievers, for each classroom. The bigger the slope residual, the more initially high-attaining pupils progress in relation to initially low achieving pupils in this classroom; the smaller the slope residual, the more this classroom contributes to more equitable rates of progress for initially high- and low-attaining pupils.



Graph 12: Classrooms ranked according to their intercept and slope residuals derived from progress model 2 in mathematics. 95% confidence intervals are applied. Intercept residuals enable comparisons between classrooms in terms of progress whereas slope residuals enable comparisons between classrooms in terms of degree of differentiation.

The top graph shows that for mathematics from a total number of 83 classrooms there are 6 classrooms (7.2%) in which pupils progressed at an above-average level, 10 classrooms (12%) in which pupils progressed significantly below average, while the majority (80.7%) of the classrooms, had average progress rates. It can be concluded that a classroom a pupil attends makes a difference to his/her educational progress, since classrooms with similar intakes in terms of initial attainment achieve very different outcomes.

The bottom graph shows that for mathematics only 2 classrooms are above average and only 3 classrooms are below average in terms of differential progress rates. It can be

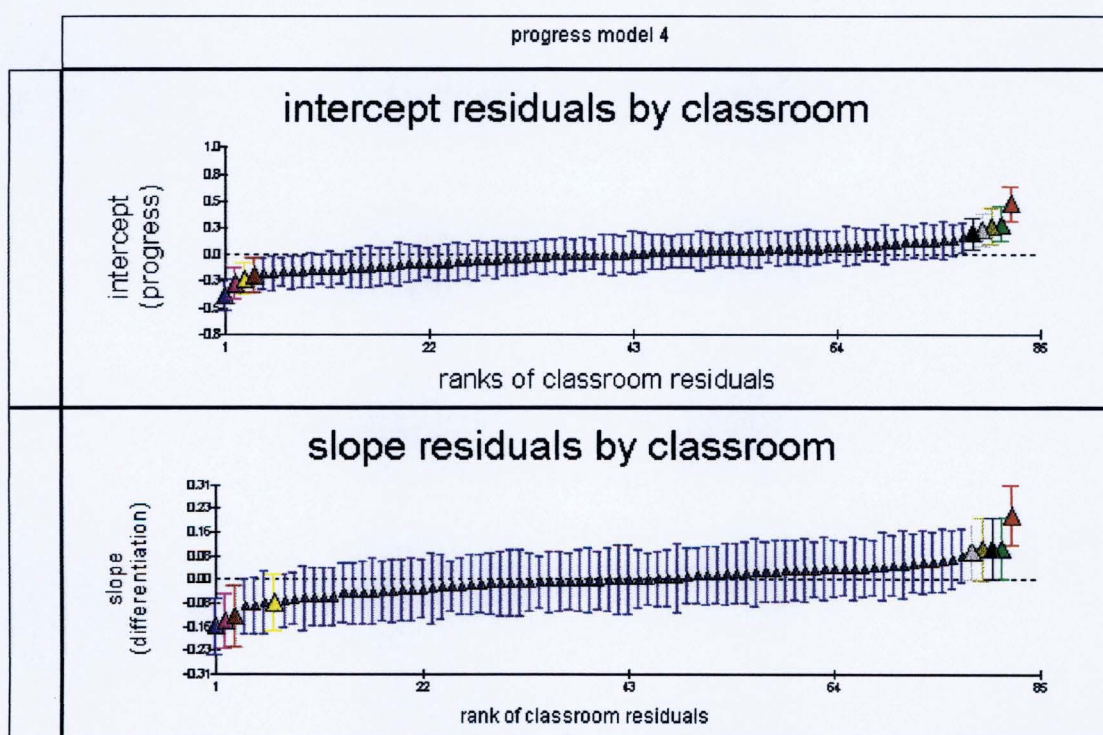
suggested that although differential progress rates exist, they are not a very pertinent characteristic of the classrooms in Piracus.

Nevertheless, intercept and slope residuals should be co-examined. All combinations are possible (high intercept and high slope, high intercept and low slope, low intercept and high slope or low intercept and low slope).

Residuals represented by the same colour in both graphs represent the same classroom. For example, in the top graph, the classroom represented in the fourth highest position has an intercept residual above average, which is represented in green and it has an average slope residual represented by the same green colour that touches the average line, as the bottom graph shows. It is evident that classrooms with above-average progress rates are average in terms of differentiation, and classrooms with below-average progress rates are average in terms of differentiation.

According to progress model 2, the confidence interval associated with the intercept residual of the classroom with the highest progress rates in mathematics included the points from 0.274 to 0.754. The 6th classroom from the top, which has the lowest progress among the above-average classrooms, had a confidence interval that included the points from 0.051 to 0.499. The classroom adjacent to it (achieving the highest from the average classrooms) has a confidence interval that includes the points from -0.013 to 0.507. A typically average classroom has a confidence interval that includes the points from -0.244 to 0.223. The classroom with the lowest progress rates has a confidence interval that includes the points from -0.575, to -0.094. The 10th classroom with below-average progress rates (the classroom with the highest progress rates among

those having scored at a below-average level) had a confidence interval that includes the points from -0.45 to -0.02. The next classroom, which is in the 11th rank from the bottom and can be characterised as average, as its confidence interval touches the average horizontal line, does not differ from the previous one so much; its confidence interval includes the points from -0.442 to 0.019. It is evident that, although we can clearly distinguish between above average, average and below-average classrooms, there are no clear dividing lines between classrooms that performed at a below-average level and classrooms with adjacent rankings which performed at an average level, as their confidence intervals slightly touch the average line. Neither are there any dividing lines between classrooms that performed at an above-average level and classrooms with adjacent rankings, which performed at an average level.



Graph13: Classrooms ranked according to the intercept residuals and to their slope residuals of the progress model 4 for language. 95% confidence intervals are applied.

In the case of language as well, the majority of classrooms performed as expected in terms of progress (at an average level), as the confidence intervals attached to classroom residuals cross the average horizontal line. The top part of graph 13 illustrating intercept residuals shows that from a total number of 83 classrooms there are 5 classrooms (6%) with above-average rates of progress, 4 classrooms (4.8%) with below-average rates of progress, while the remainder (89.1%) had average rates of progress.

The bottom part of graph 13 shows that most classrooms have slope residuals not significantly different from average. There is only one classroom with an above-average slope residual and three classrooms with below-average slope residuals, whereas for the majority of classrooms, slope residuals are roughly the same.

Next the confidence interval associated with the intercept residuals of the highest and the lower scoring classrooms were examined using a procedure similar to that employed for mathematics. Similar conclusions were reached, namely that there are no clear dividing lines between classrooms that performed at a below-average level and classrooms with adjacent rankings, which are characterised as 'average' because their confidence intervals slightly touch the average line. Nor are there any dividing lines between classrooms that performed at an above-average level and classrooms with adjacent rankings, which are characterised as average because their confidence intervals touch the average line.

The same rationale applies for all residual plots, concerning intercept and slope residuals derived from attainment models, progress models and contextual models.

3. 28. Information that classroom residuals can provide

Residuals derived from a null model can only be used to assess the relative standing of a given classroom in relation to the other classrooms that took part in the study in terms of pupils' unadjusted final attainment scores. However classroom impact in raising pupils' outcomes should not be assessed according to residuals derived from a null model, as pupils' high final attainment scores in certain classrooms may be attributed to the fact that to start with these classrooms served a 'good' cohort of high-attaining pupils.

Classroom residuals derived from absolute-attainment models assess the relative standing of a given classroom in relation to the remaining classrooms in terms of pupils' final attainment scores adjusted for individual pupils' characteristics included in the model. Conversely, classroom residuals derived from a 'progress model' show the degree to which a given classroom contributed to raise pupils' attainment; in other words they show the pooled average progress of this classroom. They estimate the additional impact of each classroom in raising pupils' final attainment scores. They can be associated with pedagogical work undertaken by the classrooms involved and can reveal pedagogical dimensions at interplay during the period of interest (the period between the two measurements). This methodological path belongs to a process-product research design though, and has not been pursued by the current study. In addition, the position of each classroom in terms of progress can be associated with the efforts and initiatives undertaken by principals, and support class teachers. Positive progress rates may be attributed to the existence of support schemes in the school. Alternatively, they may be ascribed to the sustained efforts of the mainstream and/or support teacher who collabo-

ratively and after having adopted a joint action plan, managed to deal with gaps in pupils' knowledge and raised standards. Often, enrichment initiatives at the school level adopted in the framework of the whole-day school in the afternoons, such as computer literacy, or the organisation of theatrical performances in the school, may have a positive impact on all pupils' attainment in basic skills (in literacy and numeracy). In other school cases the involvement of the wider community in school affairs might have a positive impact, thus raising attainment and producing positive progress rates for all pupils in the school.

If classrooms serve a disadvantaged pupil intake, classroom residuals derived from an adjusted progress model that controls for conditions of disadvantage are pulled towards a higher position in comparison to their position when derived from an unadjusted progress model.

In addition, it is difficult for classrooms having been positioned in the highest ranks in terms of absolute attainment, to be positioned even higher in relation to the rest, thus demonstrating positive progress rates. For highly achieving classrooms ceiling effects in their progress residuals emerge. Therefore to assess the endeavour undertaken by each classroom, residuals from both the null model and the progress model are needed. Teachers serving in initially high performing classrooms have already achieved very good results with the support of principals, management and school organisation, and hence their venture is focused on retaining this high place. On the other hand, it seems to be easier for classrooms with below-average initial attainment to show positive progress rates after having been identified; especially when this classroom is monitored carefully and provided with assistance. Examples of assistance may be teacher training,

in order that teachers may improve their teaching strategies; alternatively, support teachers can assist low achieving pupils to overcome their weaknesses. The joint impact of initiatives taken at the school level and at the classroom level can have the effect of improving pupils' attainment in low achieving schools. Classroom relative standing would improve if these classrooms demonstrated positive progress rates during the years following their identification.

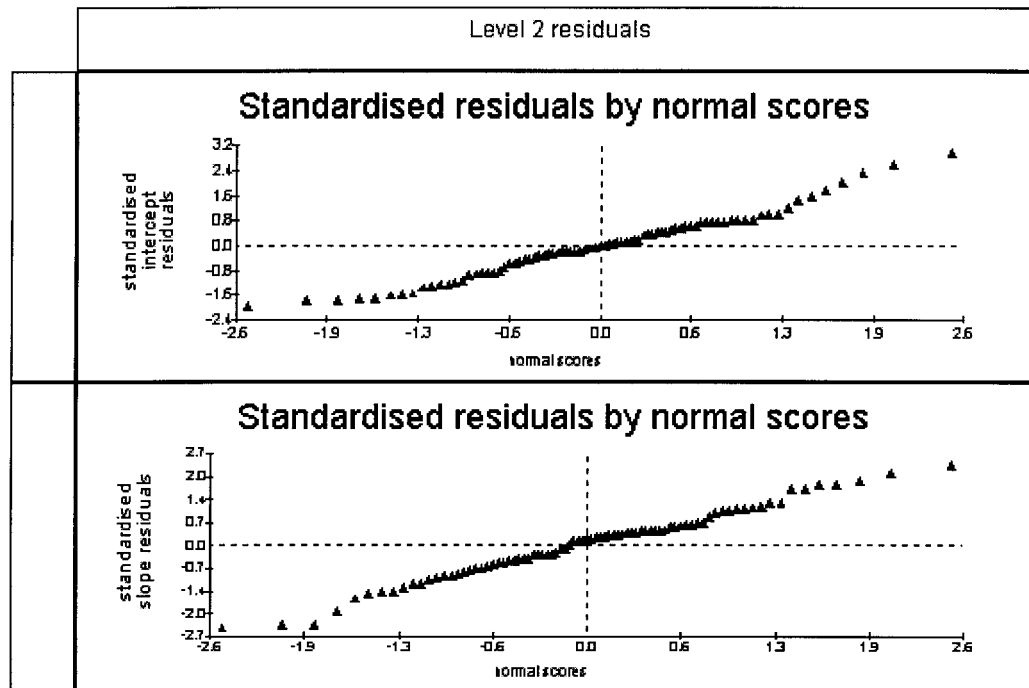
Strand (1998, p. 135) states that *'a case can be made that both raw and value added results are needed for different, although related, purposes. Raw results tell us about absolute standards of pupil attainment. They are of vital importance for individual pupils' future life chances, for monitoring attainment across the country as a whole and for identifying possible low levels of attainment by some social groups. However, pupils' raw results can say little about how well the school they have attended has contributed to their success'*. Pupils' outcomes derived from unadjusted absolute-attainment models obtained at the end of a given school year can serve to identify pupil groups which underachieve so that services tailored to their needs can be set up during the following school year. Funds can be directed to these schools in order to alleviate educational inequalities between these pupil groups. However, school and classroom residuals derived from absolute-attainment models should not be employed to hold individual schools or classrooms accountable. Such criteria are not adequate to assess how well school policies and teaching practices worked; progress criteria should be applied instead. Residuals derived from the progress model are more informative about classroom contribution in raising pupils' educational outcomes. Absolute-attainment criteria may serve to identify groups of pupils whose outcomes lagged behind the out-

comes of the majority group, so that compensatory steps are taken in the direction of equal educational opportunities. School and classroom outcomes based on absolute-attainment criteria obtained at the beginning of a school-year may identify school and classroom needs, so that remedial and support schemes can be designed and implemented. If compensatory programmes operate in a school, positive progress rates for these pupil groups in relation to the majority group would appear during subsequent school years, which eventually would cause attainment differences to be alleviated.

3. 29. Checking model assumptions

3. 29. 1. Checking the assumption of normality for level 2 residuals

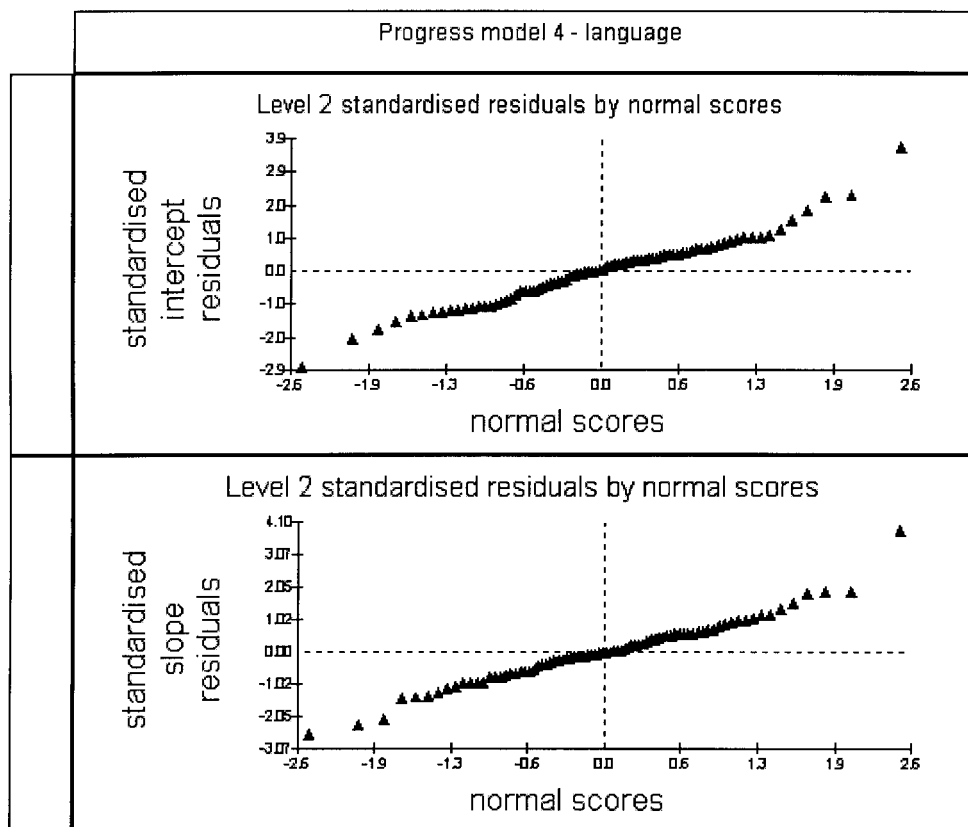
The *assumption of normality* for intercept and slope residuals at classroom level will be checked below through the graphs illustrating the ‘*standardised residuals by normal scores*’ a) derived from progress model 2 for mathematics and b) derived from progress model 4 for language.



Graph 14: Standardised residuals by normal scores for classroom intercept and slope residuals for mathematics.

b) Language

The two graphs below depict classroom intercept and slope residuals derived from progress model 4 for language.

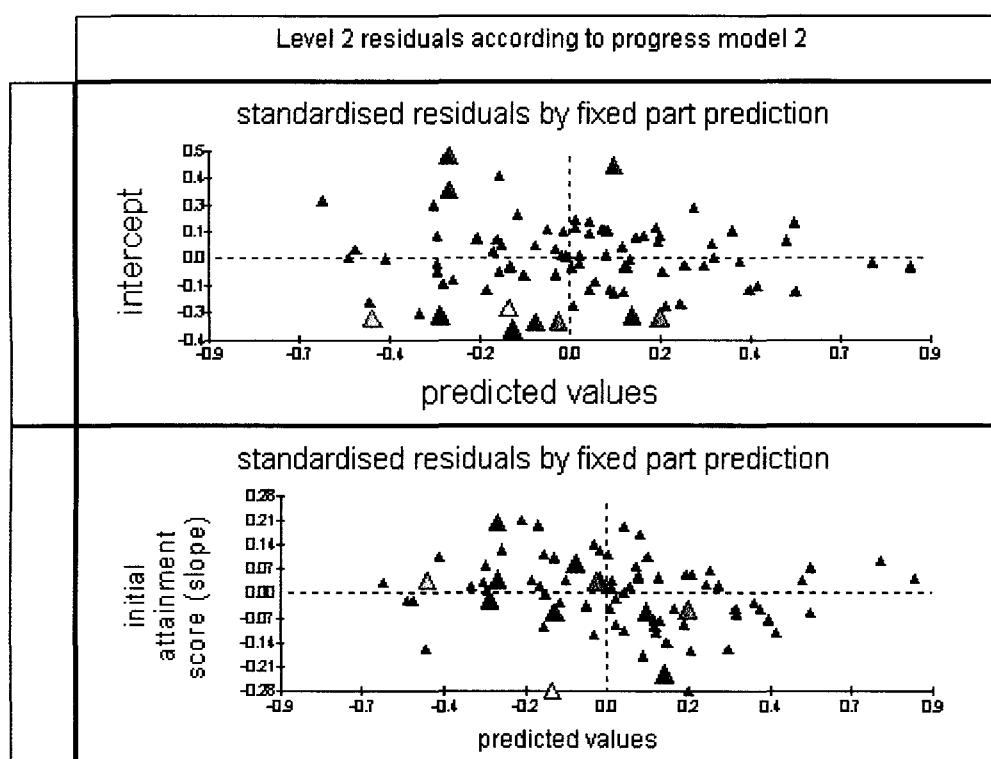


Graph 15: Level 2 standardised intercept and slope residuals and their normal scores for language

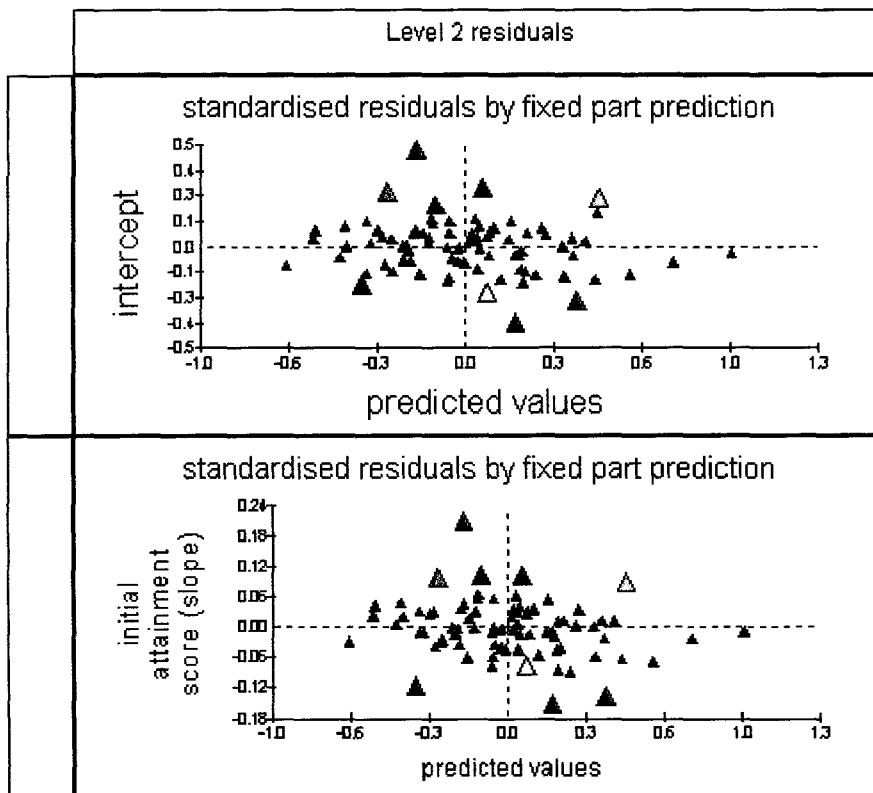
For both mathematics and language the first of the graphs shows classroom ‘*standardised intercept residuals by normal scores*’ while the second graph shows classroom ‘*standardised slope residuals by normal scores*’. It is evident that both classroom intercept and slope residuals are placed on a straight line. Therefore it can be assumed that for both mathematics and language classroom intercept and slope residuals approximately follow the normal distribution.

3. 29. 2. Checking the assumption of constant variance among level 2 residuals

The graphs below show classroom 'standardised residuals by fixed part prediction' at average initial attainment score. It can be seen that for both subjects level -2 (class-room) intercept and slope residuals have constant variance. They look more like a 'cloud' and there is no discernible pattern among them.



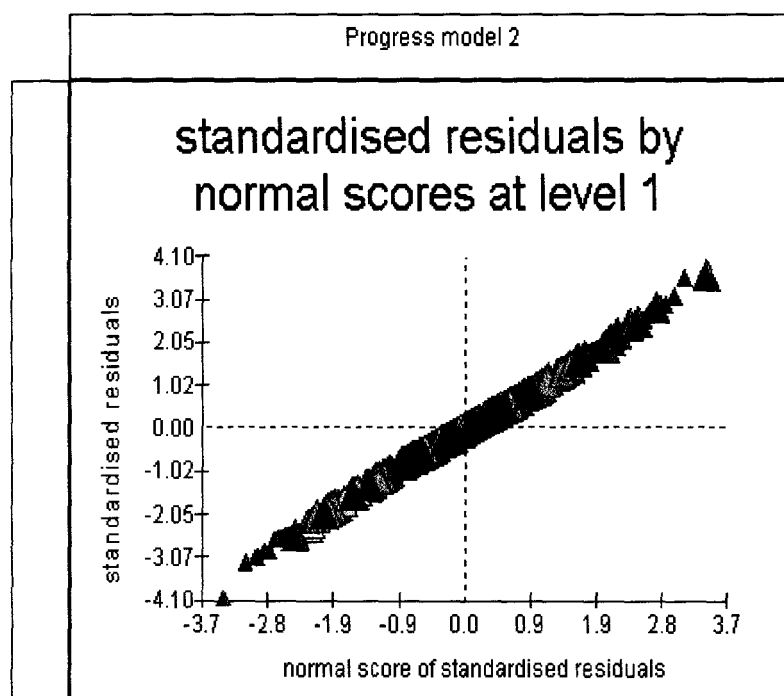
Graph 16: Standardised residuals by fixed-part predictions at level 2 at average initial attainment score for mathematics.



Graph 17: Standardised intercept and slope residuals at level 2 by fixed-part prediction for the average initial score of each class for language.

3. 29. 3. Checking the assumption of normality for level 1 residuals

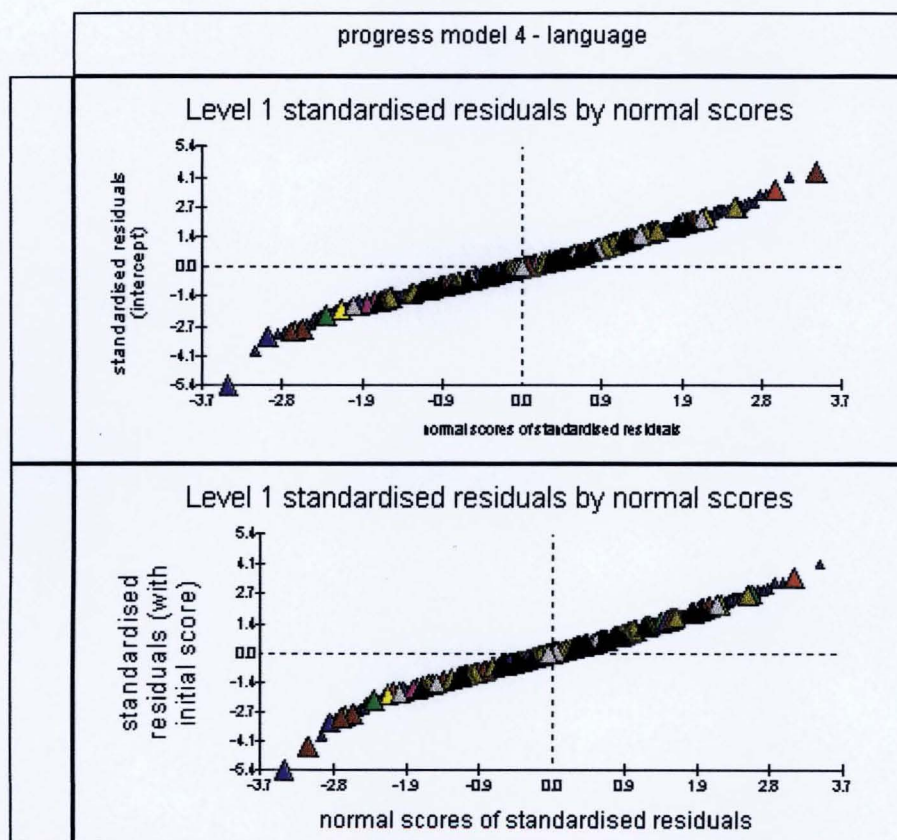
The following graph illustrating pupils' '*standardised residuals by normal scores*' shows that the level-1 (pupil level) residuals follow the normal distribution.



Graph 18: Standardised residuals by normal scores at level 1 derived from progress model 2 for mathematics

Two graphs of pupils' residuals, derived from progress model 2, are shown below.

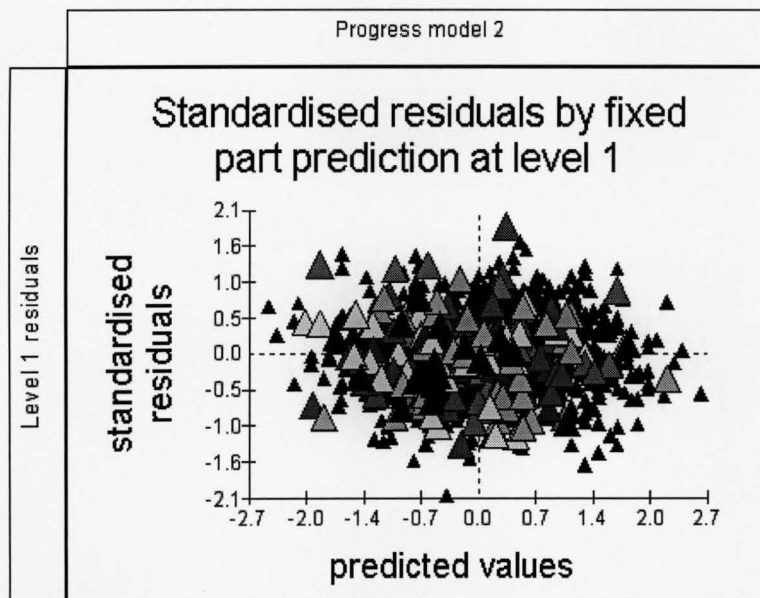
These graphs illustrate 'standardised residuals by normal scores' at pupil level. Both the graphs below show that both intercept and slope residuals are roughly positioned in a straight line. The distributions of intercept and slope residuals approximate the normal distribution; in other words the assumption of normality is satisfied.



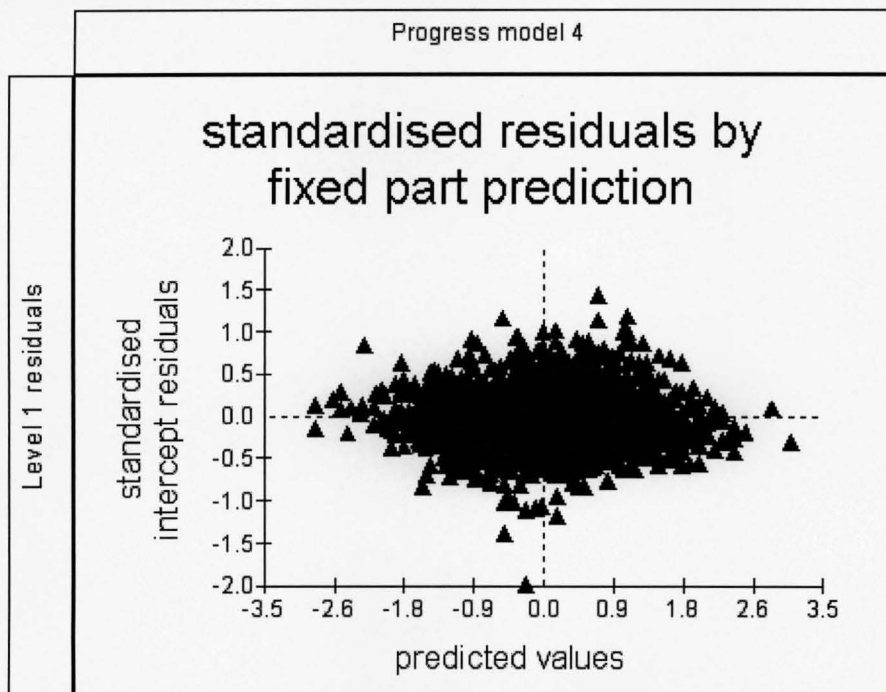
Graph 19: Standardised residuals by normal scores at level 1 for 'progress model 4' for language

3. 29. 4. Checking the assumption of constant variance for level 1 residuals

The graphs below illustrate the *standardised residuals by fixed part prediction* at pupil level for mathematics and for language.



Graph 20: Level 1 standardised residuals by fixed part prediction for mathematics



Graph 21: Level 1 standardised residuals by fixed part prediction for language

The above graphs show that level-1 residuals both for mathematics and for language appear to have constant variance. They do not show signs of heteroscedasticity, that is unequal values of variance for different values of predicted final attainment score, as no fanning in or fanning out is shown in these graphs.

3. 29. 5. Checking the linearity assumption

A quadratic term of pupils' initial attainment scores was introduced in both progress model 2 for mathematics and in progress model 4 for language in order to test whether the relationship between pupils' final attainment scores and their initial attainment scores was linear, and also to allow for possible ceiling effects related to high values of pupils' initial attainment scores. The progress models including quadratic terms are named '2c' for mathematics and '2e' for language. These models were not an improvement on 'progress model 2' for mathematics ($X^2_1 = 0.1, p=0.751$) or 'progress model 4' for language ($p=0.0783$).

3. 30. *Profiling classrooms*

The interrelationships between classroom characteristics in the setting under investigation can be studied by examining the correlation coefficients between a set of characteristics at a time. These relationships are reflected on the scatter plots between classroom intercept and slope residuals, derived either from the same model or from a range of models. An example of such an investigation would be to examine the relationship between classroom average final attainment score (classroom intercepts) by classroom degree of differentiation (slopes). Such an examination can lead to a more comprehensive picture of the characteristics of the educational system investigated. In addition it can indicate degrees of similarity between a given classroom and the rest in relation to the two characteristics investigated.

In the current analysis the *classroom intercept and slope residuals were co-examined* corresponding to:

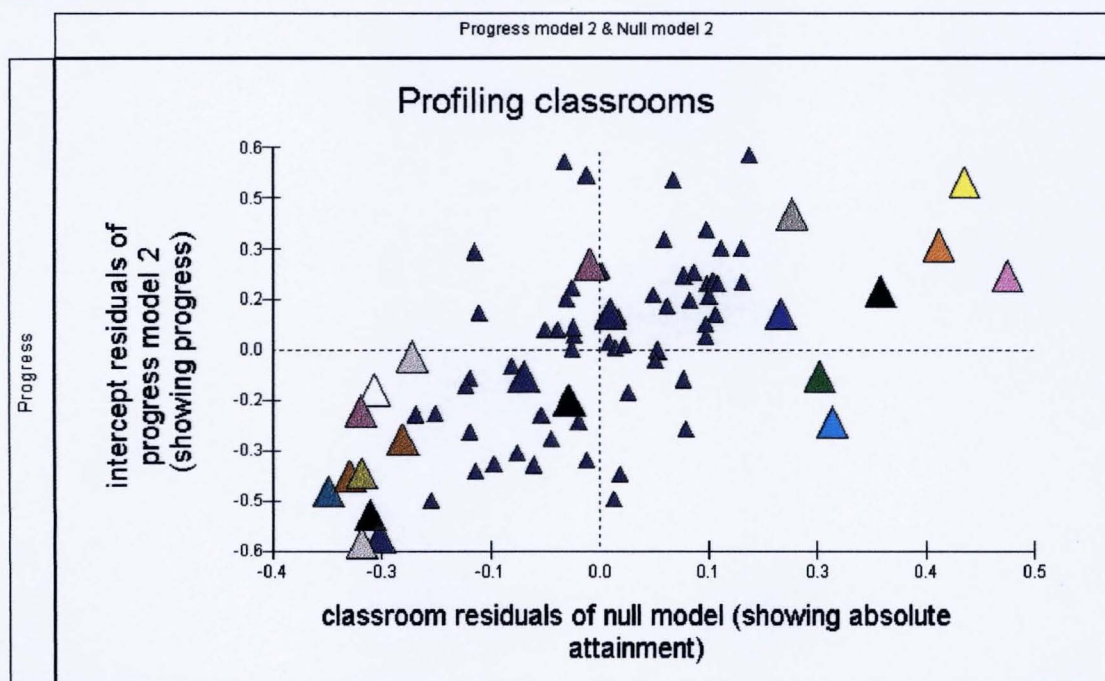
- The *null model* that shows each classroom's absolute standing in terms of unadjusted absolute attainment (intercept residuals only are derived).
- The *progress model*. The intercept residuals derived from the unadjusted progress model, (which controls only for pupils' initial attainment scores) show the relative standing of a given classroom in relation to the rest, in terms of progress.

a) Absolute attainment by progress

The graph below is a scatter-plot of classroom intercept residuals of the null-model with classroom intercept residuals of '*progress model 2*' for mathematics. Each triangle represents an individual classroom. Associated with each of these there is a standard error, which measures the uncertainty associated with sampling variation. By clicking on each of the triangles in the MLWin software graph, the analyst can identify individual classrooms. According to this scatter plot, classrooms can be classified into four categories:

- a) Classrooms with average or above average absolute attainment intercept residuals according to null model and average or above average progress are represented in the top right part of the scatter plot.
- b) Classrooms with average or above average absolute attainment and average or below average progress are represented in the bottom right part of the scatter-plot.
- c) Classrooms with average or below average absolute attainment and average or above average progress are represented in the top left part of the scatter plot.
- d) Classrooms with average or below average progress and average or below-average absolute attainment are represented on the bottom left part of this plot.

Graph 22: Scatter plot of classroom intercept residuals according to progress model 2 (showing progress) with intercept residuals according to null model 2 (showing absolute attainment) for mathematics.



Classrooms that score high in terms of absolute attainment tend to score high in terms of progress. The positive association between the intercept residuals of the null model (revealing a classroom's unadjusted performance) and the intercept residuals of progress model 2 (revealing a classroom's progress) shown in the above graph is also shown by the moderately high correlation coefficient between them ($r = 0.63$, $p < 0.001$). Many classrooms that have average or above-average intercept residuals according to the null model have also average or above-average intercept residuals according to 'progress model 2'.

For mathematics there are classrooms that are ranked as above average in terms of absolute attainment but slightly below average in terms of progress. The classroom represented in a light blue triangle in the previous graph is such a case. Pupils en-

rolled in this classroom had above-average unadjusted performance at entry and thus the classroom has not shown a lot of progress, but still retained its above-average status at the end of the school year. Ceiling effects make it difficult for classrooms with an above-average initial attainment score to demonstrate also high progress rates. If a classroom has scored high at the beginning, there is little room for improvement. If a classroom has scored 90 on a 100-points scale, the 10 extra units are more difficult to obtain than they would be for a classroom having scored 50 on the same scale.

In the case of language, the correlation coefficient between intercept residuals according to null model 2 and the intercept residuals according to 'progress model 4' is 0.386 ($p < 0.001$). This correlation implies that for language, there is a weak positive relationship between classrooms' unadjusted performance and their progress rates.

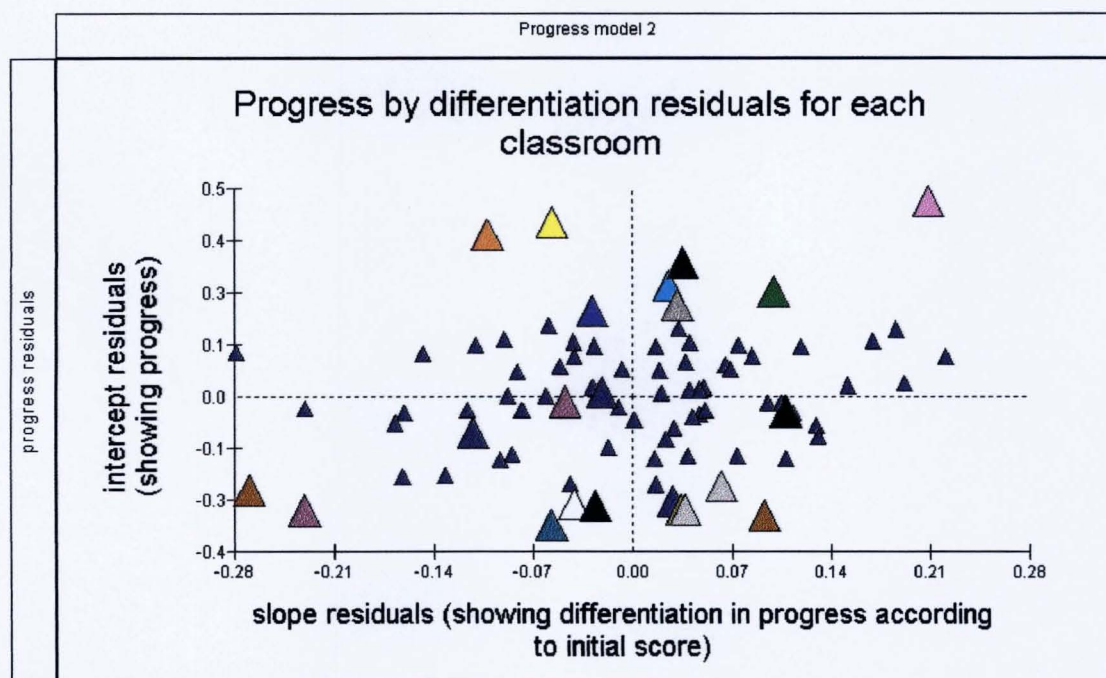
b) Absolute attainment by degree of differentiation

There is no association between classrooms' unadjusted performance level at the end of the year and the extent to which classrooms contribute differently to the progress of high achievers compared to the progress of low achievers. The correlation coefficient between intercept residuals of null model 2 and the slope residuals of progress model 2 is -0.04 , in practical terms zero.

In the case of language, the correlation coefficient between intercept residuals according to null model 2 and the slope residuals according to progress model 4 is 0.327 ($p=0.003$), which implies that there is a weak positive relationship between classrooms' unadjusted performance and the extent to which classrooms produce different rates of progress for high achievers from those for low achievers.

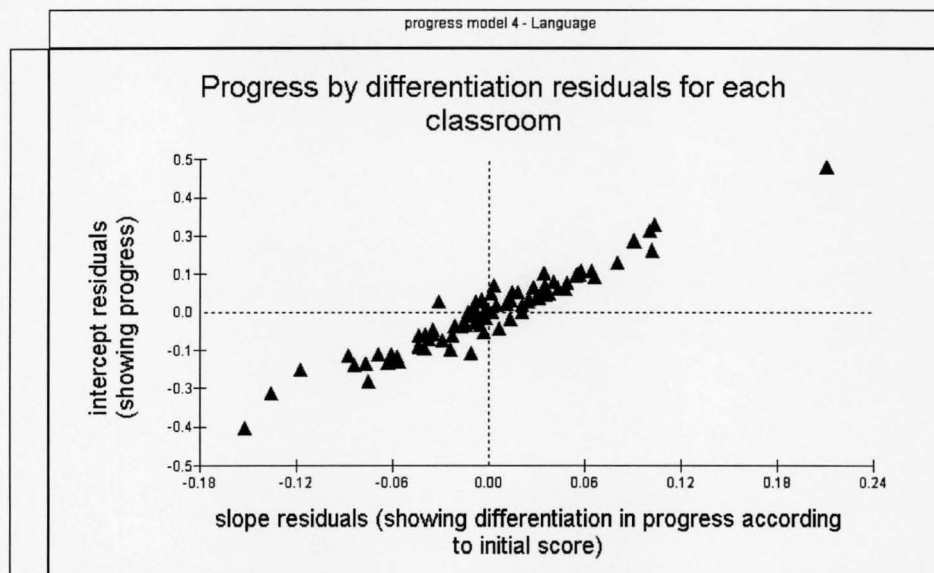
c) Progress by degree of differentiation

The two scatter-plots below plot classroom intercept-residuals with classroom slope-residuals according to progress model 2 for mathematics and progress model 4 for language. Each triangle represents an individual classroom. These scatter plots explore whether there is a relationship between the progress made by each classroom and the extent to which the classroom has produced differential progress rates for initially high-attaining pupils in relation to low achieving pupils.



Graph 23: Scatter plot of classroom intercept by slope residuals according to progress model 2 for mathematics.

The above plot shows no association between classrooms' progress rates and the extent to which classrooms produced different progress rates for initially high achieving pupils compared with those for initially low achieving pupils. The correlation coefficient between intercept and slope residuals derived from progress model 2 is very small and not significant: ($r = 0.187$, $p=0.09$).



Graph 24: Scatter plot of classroom intercept by slope residuals according to progress model 4 for language.

For language, the correlation between intercept and slope residuals derived from progress model 4 is 0.962 ($p < 0.001$), which implies that in classrooms where pupils progress well high achieving pupils progress more than low achieving pupils. The above scatter plot shows that there is a strong linear relationship between classroom intercept and slope residuals. Classrooms with a big intercept tend also to have a big slope. This finding can be contrasted with the absence of relationship between classroom intercept and slope residuals for mathematics.

3. 31. Is it legitimate to make comparisons between classrooms?

Researchers and data analysts should be unwilling to undertake comparisons between individual classrooms, especially when those comparisons are undertaken for accountability purposes. Goldstein (1997) highlighted the point that *'institutional comparisons have to be based upon suitable adjustments for intake achievement and other relevant factors'*. However, even after having adjusted for individual pupils' characteristics that can be considered as givens, there is always the risk of not having adequately controlled for all intervening variables that require an extra effort on the teachers' part as they may have an impact on pupils' progress in the class. *'It is quite possible that even the best available models lack key 'contextualising' variables thus emphasising the additional caution that needs to be taken with interpretations'* (Goldstein et al., 2000, p. 11). Rowe (2000, p. 81) also suggested that *'there is always the difficulty that any statistical model will fail to incorporate all the appropriate adjustments'*.

Even after controlling for pupils' initial attainment scores and for individual pupils' characteristics there may still be differences between classrooms in terms of progress, which although out of the realm of classrooms or schools, may not have been controlled for. Controlling for these intervening variables would have enabled *'fairer'* comparisons between classrooms to be undertaken. Factors that can be considered as givens are not related to characteristics that stem from successful pedagogical approaches; the latter are not characteristics that a model should control for

in order to enable fairer classroom comparisons to be undertaken. By contrast, factors out of the realm of the school can be considered as givens. Such factors are pupils' social class, ethnicity, their competence in their target language, gender, etc. These characteristics can be thought of as factors facilitating or inhibiting pedagogical practices. Schools / classrooms have to take extra steps and initiatives to alleviate the effects associated with conditions of disadvantage. Such steps might be the adoption of support-teaching programmes, amendments in mainstream class programme and curricula or the employment of more staff.

3. 32. Social class and Mother's education

The Goldthorpe social class variable was introduced in an absolute attainment model, in mathematics, to identify the impact that belonging to each social class category has on pupils' final attainment scores. Given that the highest social class category has been used as the base category, the remaining eight categories were introduced as dummy variables in the models (with no other explanatory variables). The resulting social class coefficients (*with their standard errors in brackets*) derived from the absolute attainment model for mathematics are shown below:

Table 3. 14: Fixed part coefficients of social class derived from absolute attainment model 11 for mathematics according to the 9-category Goldthorpe schema

Goldthorpe class designation	Description	Fixed part coefficients
1	High grade professionals, administrators and officials; managers in large industrial establishments; large proprietors	Base category
2	Lower grade professionals, administrators and officials; higher-grade technicians; managers in small industrial establishments; supervisors of non-manual employees	0.205 (0.170)
3	Routine non-manual employees, higher grade (administration and commerce)	-0.068 (0.170)
4	Routine non manual employees, lower grade (sales and services)	-0.274 (0.160)
5	Small proprietors, artisans, etc., with employees	-0.177 (0.199)
6	Small proprietors, artisans, etc., without employees	-0.445 (0.169)
7	Lower-grade technicians; supervisors of manual workers	-0.372 (0.192)
8	Skilled manual workers	-0.549 (0.161)
9	Semi- and unskilled manual workers (not in agriculture, etc.)	-0.665 (0.168)

It is apparent that only four coefficients are significant or near significance level. These coefficients correspond to adjacent social class categories. These are the coefficients of categories 6, 7, 8, and 9.

These dummy variables were, at a subsequent phase, re-introduced in a model employing a dichotomous social class classification schema. According to this *dichotomous social class classification schema*, the last four coefficients of the 9-category social class schema test were grouped together and their impact was tested against the remaining 5 higher categories. In this way, social class categories 6, 7, 8 and 9 were grouped together and their impact was contrasted with the impact of social class categories 1, 2, 3, 4 and 5.

The variable representing mother's education has seven categories and is defined as the highest educational level completed by a pupil's mother. Initially the variable was introduced on its own in absolute attainment models, for mathematics and for language, so as to identify the impact that the mother having completed an additional educational level has on pupils' final attainment scores. The coefficients derived from these absolute attainment models are shown in the following table 3.15:

Table 3. 15: Classification scheme for mother's educational level

Coding	Mother's educational level:	Coefficients of 'mother's education' categorical variable derived from the abso- lute attainment model for mathematics ($p < 0.001$). (Standard errors are in brackets).	Coefficients of 'mother's education' categorical variable derived from the abso- lute attainment model for language ($p < 0.001$). (Standard errors are in brackets).
0	She is completely illiterate	<i>Base category</i>	<i>Base category</i>
1	She was enrolled in some grades in primary school.	0.068 (1.019)	1.485 (0.987)
2	She has completed primary school.	0.425 (0.932)	2.014 (0.915)
3	She has completed compulsory education (she has completed <i>Gymnasio</i> that corresponds to nine years of schooling)	0.480 (0.932)	2.055 (0.916)
4	She has successfully completed secondary education (she has completed <i>Lykeio</i> that corre- sponds to 12 years of schooling).	0.782 (0.932)	2.392 (0.915)
5	She is a Tertiary education gradu- ate but her degree is not equiva- lent to a University degree.	0.880 (0.938)	2.428 (0.921)
6	She is a Higher education gradu- ate (she has a University degree).	1.163 (0.933)	2.744 (0.917)

This classification scheme is the same as that employed by the National Statistical Service of Greece (NSSG).

Next, the initial four categories (0-3) were collapsed to form the base category in a new dichotomous 'mother's education' variable. The base category of this variable (which was coded with zero) was comprised of pupils whose mothers either *completed or did not complete compulsory education*, which corresponds to nine years of schooling. The

remaining three categories (4-6) of the detailed mother's education variable were collapsed to form the '*higher than compulsory education*' category, which is assigned the value of 1. In this category were assigned pupils whose mothers had completed more than the nine years of compulsory education.

3. 33. Differential effectiveness

The impact that individual pupils' characteristics have (such as initial attainment score, ethnicity, social class or gender) on pupils' progress may be different for different classrooms. The occurrence of different impact for different pupils' characteristics on pupils' progress rates in different classrooms is known as differential classroom effects. Differential effects have been previously identified in relation to pupils' initial attainment score for both subjects.

Classroom impact in terms of progress may vary systematically for different pupil groups. The hypothesis that classrooms may be differentially effective for different pupil groups is tested by examining whether random coefficients at the classroom level (level-2) can be assigned to variables representing individual pupils' characteristics. Through the differential effects hypothesis one can examine whether classrooms differ to the extent that they enable pupils with different individual characteristics to perform optimally. Therefore, the current study has tested whether the impact of all variables representing individual pupils' characteristics such as ethnicity, gender, social class, friendships, etc. on pupils' progress is the same for all classrooms investigated, or different for different classrooms. If differential effects exist, then the predicted classroom

lines would have different slopes associated with variables with random coefficients, instead of being parallel.

3. 34. Models accounting for individual pupils' characteristics

Based on progress model 2 for mathematics and on progress model 4 for language a series of progress models has been built, estimating the impact of individual pupils' characteristics on pupils' progress during the school year examined. Adjusting for individual pupils' data concerning pupils' socio-economic status and ethnicity along with their initial attainment score is necessary to ensure that classroom comparisons do not unfairly reflect on classrooms serving a majority of disadvantaged pupils. Such a progress model that additionally adjusts for individual pupils' characteristics is progress model 3d2 for mathematics, shown in figure 9 below. This progress model along with initial attainment score additionally accounts for pupils' ethnic status, their social class membership and the interaction term 'ethnic status by initial attainment score'. This model appears also in table 4. 9: Significant interaction effects in progress models for mathematics and language are in section 4. 3. 10.

Figure 9: Progress model 3d2 adjusting for pupils' initial score and for pupils' social class, ethnicity (being of foreign ethnic background), and for the interaction term 'being of a foreign/repatriated ethnic origin by initial attainment score'

Equations

$$\text{enwt_scr}_{ij} \sim N(XB, \Omega)$$

$$\text{enwt_scr}_{ij} = \beta_{0ij}\text{constant} + \beta_{1j}\text{nt_scor2}_{ij} + -0.066(0.032)\text{p_cls}(1_lower)_{ij} +$$

$$-0.070(0.048)\text{rec_ethnic1}(1\text{-foreign})_{ij} +$$

$$0.098(0.047)\text{foreigner.nt_scor2}_{ij}$$

$$\beta_{0ij} = 0.048(0.034) + u_{0j} + e_{0ij}$$

$$\beta_{1j} = 0.789(0.024) + u_{1j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} 0.049(0.011) & \\ & 0.004(0.006) \end{bmatrix}$$

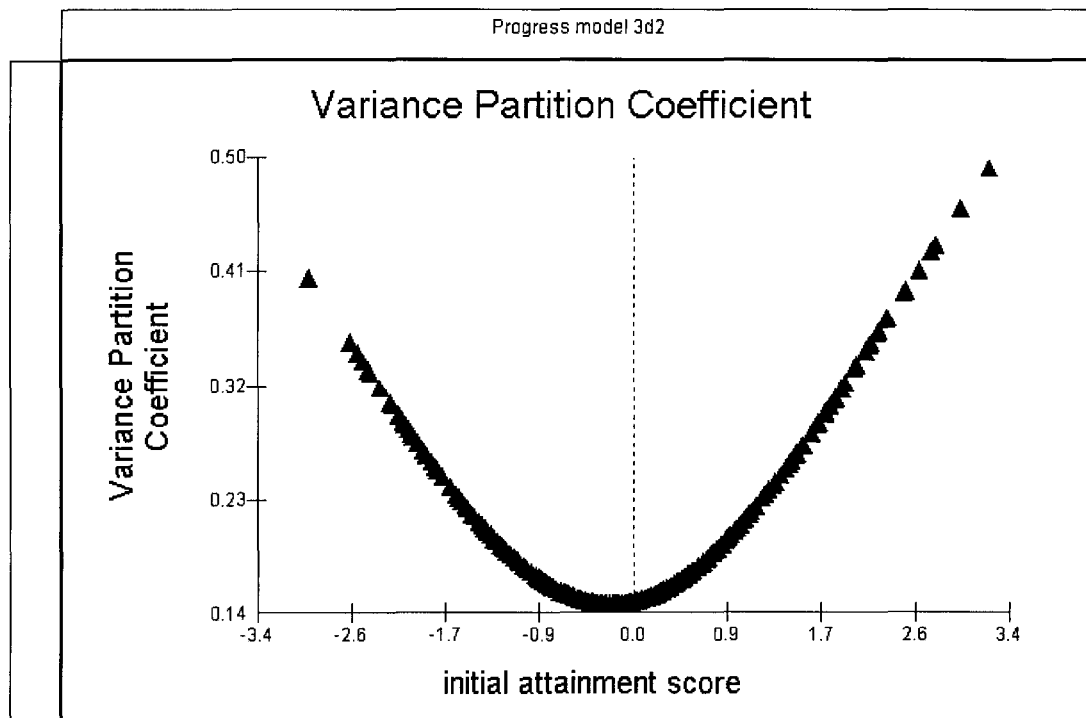
$$e_{0ij} \sim N(0, \Omega_e) : \Omega_e = [0.279(0.011)]$$

-2*loglikelihood(IGLS) = 2239.461(1333 of 1858 cases in use)

Fonts Subs Name + - Add Term Estimates Nonlinear ? Help Clear

The model including the interaction term 'foreign by initial score' (3d2) is a significant improvement on the model with ethnicity and social class (3d) at $p=0.04$, as the deviance test showed ($X^2_1=4.2$). The above p-value comes from a chi-square distribution with 1 degree of freedom, as this model employed one extra parameter compared to the previous model, for the variable representing the interaction term 'foreign by initial score'. This positive interaction term implies that the higher foreign/repatriated pupils attained at the beginning of the school year, the more progress they made.

In graph 25 below the relationship between the VPC and pupils' initial attainment scores is shown for progress model 3d2. Similar to the variance function, the value of VPC reaches a minimum of 14.7 % for initial score values of approximately zero and increases for values of initial scores greater than or less than zero.



Graph 25: Variance Partition Coefficient by initial attainment score for mathematics according to progress model 3d2 that accounts for individual pupils' characteristics.

For example, for an initial attainment score one normal score unit above average, the VPC takes the value 21.2% and for two normal score units above average it takes the value 33.4%. The values of VPC derived from the progress model incorporating individual pupils' characteristics (3d2) approximate those VPC values obtained from progress model 2 (20.7% and 39.9% respectively).

At average initial attainment score the variance reduction at level 2 of this model in relation to progress model 2 is 11.3 %. At the pupil level the variance reduction does not depend on pupils' prior attainment and is 1% only.

3. 35. Predictions according to the above 'individual characteristics model 3d2'

To investigate the extent to which classrooms influence pupils' final attainment score, predictions are made for pupils belonging to certain pupil groups according to the above progress model 3d2. According to Teddlie et al. (2000, p. 72), by including background and prior achievement variables in multilevel models measuring the impact of school (or of the classroom) researchers control for those variables considered as threats to the internal validity of the studies.

The predicted end score for pupils enrolled in classroom j according to this model is:

$$y_{ij} = \beta_0 + \beta_{1j}x_{ij} + \beta_2x_{2ij} + \beta_3x_{3ij} + \beta_4x_{3ij}x_{1ij} + u_{0j}$$

In which:

$$\beta_{1j} = \beta_1 + u_{1j}$$

Shows that the coefficient of initial score varies between classrooms,

β_2 is the coefficient of social class

β_3 is the coefficient of ethnicity and

β_4 is the coefficient of the interaction term ‘ethnicity by initial attainment score’.

If the values of the coefficients derived from progress model 3d2 are placed in the first equation, the following equation is derived:

$$y_{ij} = 0.048 + (0.789 + u_{1j})x_{1ij} - 0.066x_{2ij} - 0.070x_{3ij} + 0.098 \cdot x_{1ij} \cdot x_{3ij} + u_{0j}$$

This equation shows that every extra unit of initial score matters more for foreign pupils than for Greek pupils in terms of progress. For foreign pupils the above equation becomes:

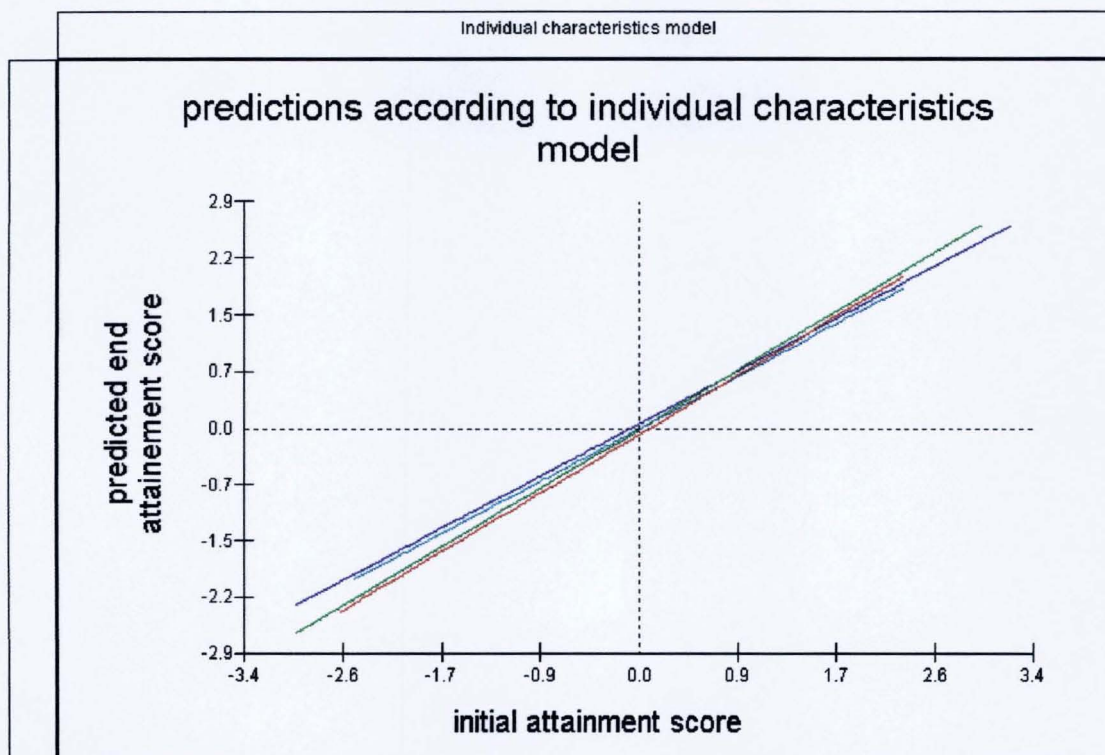
$$y_{ij} = 0.048 + (0.887 + u_{1j})x_{1ij} - 0.066x_{2ij} - 0.070 \cdot 1 + u_{0j}$$

That is to say:

$$y_{ij} = -0.022 + (0.887 + u_{1j})x_{1ij} - 0.066x_{2ij} + u_{0j}$$

While for Greek pupils, the above equation becomes:

$$y_{ij} = 0.048 + (0.789 + u_{1j})x_{1ij} - 0.066x_{2ij} + u_{0j}$$



Graph 26: Average predicted lines according to progress model 3d2 adjusting for individual pupils' characteristics for different categories of pupils for mathematics.

The four lines shown in the above graph represent the following pupil groups: The blue line represents Greek pupils from higher social class background, the light blue line Greek pupils from lower social class background, the green line foreign pupils from higher social class background, and the red line foreign pupils from lower social class background.

The above graph shows that there is some discrepancy between these groups of pupils' final attainment scores for pupils who initially scored less than average (less than zero in normal score units), with the lowest achieving groups being foreign pupils from lower social class background and foreign pupils from higher social class background. On the other hand, the final attainment scores for initially high achieving Greek and foreign pupils from different social classes hardly differ.

3. 36. *Contextual Models*

In addition to variables referring to pupils' characteristics, there may be contextual or compositional factors in the classroom setting that influence the progress of all pupils enrolled in the same class. Compositional or contextual variables are considered as level-2 variables. Accounting for the impact of these background factors would assist in better estimating classroom effects and allow fairer comparisons to be made between classrooms. School Effectiveness studies have frequently identified school-composition or school-mix effects (Nash, 2001). Willms and Raudenbush (1989) reported school compositional effects related to pupils' socio-economic status. Willms (1992) pointed out that school intake composition could influence pupils' outcomes over and above the effects of pupil individual performance at entry and social class. In the current study every individual pupils' characteristic was aggregated at the classroom level and their classroom means and the standard deviations were considered as contextual variables. Examples of compositional characteristics for each class are '*pupils' average performance*', the '*variance among pupils' initial scores*', or '*the percentage of high achievers in the class*'. '*Class size*' can be considered as a contextual variable because it is constructed independently of pupils' characteristics.

The following table (3. 16) provides summary statistics for classroom compositional and contextual characteristics.

Table 3. 16: Descriptive statistics concerning compositional and contextual characteristics of classrooms

<i>Classroom contextual characteristics:</i>	<i>Mathematics</i>		<i>Language</i>	
	<i>Mean</i>	<i>Standard Deviation</i>	<i>Mean</i>	<i>Standard Deviation</i>
Average initial attainment score	-0.002	0.4	-0.002	0.32
Variance of initial attainment score	0.9	0.2	0.9	0.15
Percentage of low achieving pupils	0.25	12.6	0.25	0.1
Percentage of high achieving pupils	0.25	0.14	0.25	0.14
Number of low achievers	0.26	0.44	0.23	0.4
Number of high achievers	0.26	0.44	0.24	0.4
Percentage of foreign pupils who are low achievers	0.05	0.05	0.07	0.26
Percentage of foreign pupils	0.14	0.1	0.14	0.1
Number of foreign pupils	2.9	1.9	2.9	1.9
Percentage of pupils from lower social class	0.5	0.2	0.5	0.2
Average number of absences in the class	4.4	1.5	4.4	1.5
Average educational level of pupils' mothers	3.8	0.5	3.8	0.5
Average amount of weekly homework	12.3	2.7	12.4	9.5

The impact of compositional or contextual factors on pupils' progress is considered to be over and above the impact of other individual characteristics already included in the progress model. In the absence of cross-level interaction terms, compositional or contextual effects have the same impact on the performance of every pupil in a given class. More general justifications for including such variables in a model are that '*higher lev-*

els can be thought of as overt measures creating effectiveness enhancing conditions at lower levels. In addition, higher-level conditions may serve as buffers to protect efficiency – enhancing conditions at lower levels’ (Scheerens and Bosker, 1997, p. 60).

Furthermore, compositional or contextual factors may interact with pupils’ characteristics. Such interaction terms are considered to include variables from two levels (e.g. from classroom level and from pupil level) and they are known as ‘*cross-level interaction’ terms*. For example, a question that can be answered through employing cross-level interaction terms is the following: ‘*Do low performing pupils do better when they are educated in a classroom with a majority of high achieving pupils or do they do worse?’*

When a slope coefficient β_{1j} (e.g. initial score coefficient) is estimated from a level-2 variable, a cross level interaction effect emerges. If this model is fitted, different slopes for the initial score variable are obtained, depending on the percentage of pupils from lower social class in each classroom.

Snijders and Bosker (1999, p. 73) explain cross-level interaction effects in the following manner:

If in a basic model:

$$Y_{ij} = \beta_{0j} + \beta_{1j}x_{ij} + e_{ij}$$

the group-dependent regression coefficients β_{0j} and β_{1j} are derived from level-two variables z ,

$$\beta_{0j} = \gamma_{00} + \gamma_{01}z_1 + u_{0j}$$

and

$$\beta_{1j} = \gamma_{10} + \gamma_{11}z_1 + u_{1j}$$

then the above two equations are introduced into the basic model:

$$Y_{ij} = (\gamma_{00} + \gamma_{01}z_j + U_{0j}) + (\gamma_{10} + \gamma_{11}z_j + u_{1j})x_{ij} + e_{ij}$$

By re-arranging terms the following equation is derived:

$$Y_{ij} = (\gamma_{00} + \gamma_{01}z_j) + (\gamma_{10} + \gamma_{11}z_j) \cdot x_{ij} + (u_{0j} + u_{1j}x_{ij} + e_{ij})$$

To quote Snijders and Bosker (1999, p. 74):

'In the last equation we see that explaining the intercept β_{0j} by a level two variable Z leads to a main effect of z , while explaining the coefficient β_{1j} of x by a level-two variable z leads to a product interaction effect of x and z . Such an interaction between a level - one and a level-two variable is called a cross-level interaction'.

Contextual variables can be considered as level-2 variables because they represent the impact of classroom characteristics on pupils' progress. If a cross-level interaction is significant, different classroom slopes associated with initial attainment score are obtained. These slopes depend on the value of the contextual variable for each individual classroom.

Otherwise, if a contextual variable is included without a cross-level interaction term, the inclusion of the contextual variable into the multilevel equation would imply that intercept residuals of classroom lines would move upwards or downwards, depending

on the sign of the contextual variable. The inclusion of a contextual variable would move classroom lines to a higher or lower position.

For language, a contextual model adjusting for the percentage of pupils who are under-achieving in the classroom is contextual model cont3, shown in figure 10 and in table 4.11.

This model can be described with the following equation:

$$y_{ij} = \beta_0 + \beta_1 x_{1ij} + \beta_2 x_{2ij} + \beta_3 x_{1ij} x_{2ij} + \beta_4 x_{4j} + (u_{0j} + e_{0ij} + e_{1ij} x_{1ij}),$$

The intercept for a classroom j β_{0j} is derived from adding the average intercept β_0 with each classroom's intercept residual u_{0j} , which takes a different value for each classroom:

$$\beta_{0j} = \beta_0 + u_{0j}$$

and similarly the slope for a classroom j β_{1j} of x_{1ij} is derived from adding the average slope β_1 to each classroom's slope residual u_{1j} :

$$\beta_{1j} = \beta_1 + u_{1j}$$

The contextual term $\beta_4 x_{4j}$ implies that the performance of all pupils in every classroom j is increased or reduced by a factor depending on the percentage of low achieving pupils in the class.

where:

$$\text{var}(e_{0ij}) = \sigma_{e0}^2, \text{var}(e_{1ij}) = \sigma_{e1}^2, \text{cov}(e_{0ij}, e_{1ij}) = \sigma_{e01}$$

y_{ij} is pupils' final attainment score

x_{1ij} is pupils' initial attainment score.

x_{2ij} is pupils' social class (dichotomous variable), and

x_{4j} is the percentage of low achieving pupils in the class.

In this model the variance at both the classroom level and the pupil level is modelled as a quadratic function of pupils' initial attainment score, which are similarly defined as the quadratic variance function of progress model 4.

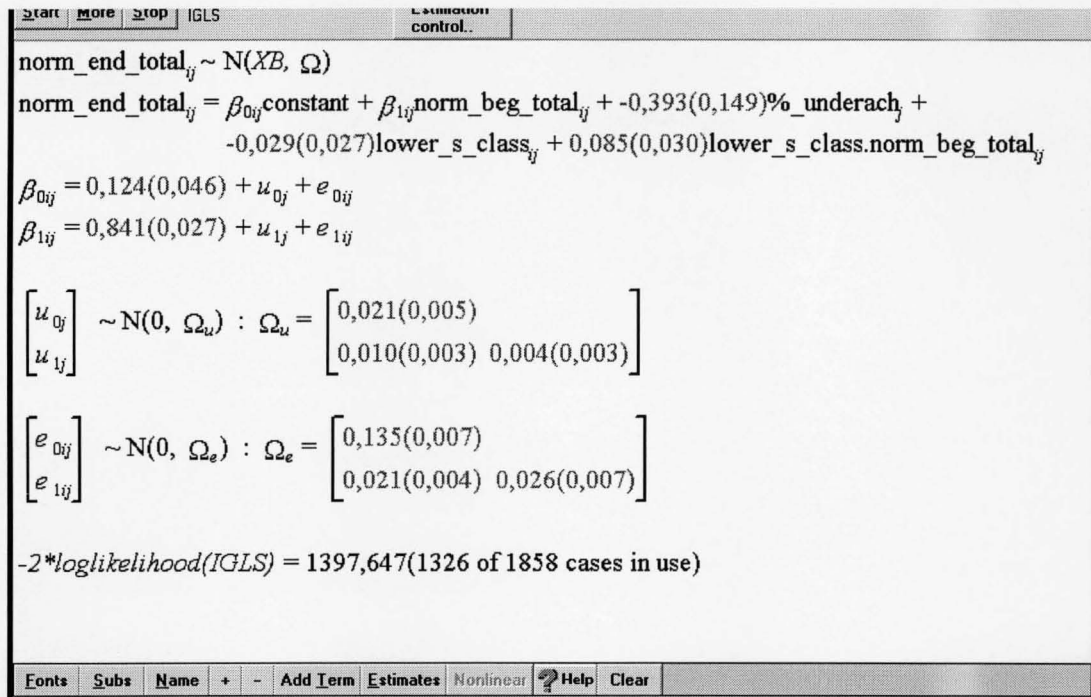
The variance at the classroom level is:

$$\text{var}(u_{0j} + u_{1j}x_{1ij}) = \sigma_{u0}^2 + 2\sigma_{u01}x_{1ij} + \sigma_{u1}^2x_{1ij}^2$$

While the variance at the pupil level is:

$$\text{var}(e_{0j} + e_{1j}x_{1ij}) = \sigma_{e0}^2 + 2\sigma_{e01}x_{1ij} + \sigma_{e1}^2x_{1ij}^2$$

Figure 10: Contextual model cont3 for language, adjusting for the percentage of pupils who are underachieving in the classroom along with initial attainment score, social class, and interaction term 'social class by initial attainment score'.



The above model measures the impact of a contextual variable on pupils' final attainment scores. Contextual model cont3 is an improvement on the adjusted progress model 9a shown in table 4. 9. that adjusts for initial attainment score, social class and the interaction of social class by initial attainment score at p=0.011. This contextual model adjusts for the previous variables and also for the 'percentage of low achieving pupils in the class'. Pupils with an initial attainment score equal to or less than -0.6214 normal score units (less than or equal to the 27 percentile of the total distribution of initial attainment score) have been defined as low achievers. The above model shows that the percentage of pupils with low initial attainment in a given class impacts negatively on the progress of all pupils enrolled in this class.

At average initial attainment score the VPC becomes:

$$VPC = 13.4\%$$

Comparison of the classroom residuals' rankings derived from contextual model cont3 with those derived from progress model 4 shows interesting differences. The reader can compare the 2nd appendix, in the second part of Table 2_6, which shows the rankings of classrooms that scored at a below-average level. One example is the 2nd classroom in school named H by the researcher for confidentiality reasons, which is shown as the 4th from the bottom and has performed at a below-average level according to a series of models (according to progress model 4, progress model 4d, null model 2 and absolute attainment model 8f). However, according to the contextual model cont3 this classroom's residual has moved to a higher position (11th from the bottom) not significantly different from the average, as shown by the confidence interval associated with this classroom. In this classroom a higher percentage of underachieving pupils were enrolled in relation to other classrooms. This classroom's relative ranking has moved upwards in the contextual model as this controls for a variable measuring a similar quality to what the dependent variable measures, that is educational attainment. The coefficient of this contextual variable reveals the impact of the presence of low achieving pupils on the average progress of all pupils in a given classroom.

Comparisons between classroom residuals derived from contextual model cont3 no longer reveal the average progress achieved by all pupils in a given classroom, but instead reveal progress adjusted for pupils' social class and the contextual variable, as classroom comparisons are based upon the unexplained variation at the classroom level. Variation associated with the presence of low achieving pupils in classrooms is removed from the random part of the model and consequently from classroom residuals.

Therefore the remaining variation at level 2 can be associated with the presence of high achieving and average achieving pupils in classrooms only. Residuals from classrooms with a high percentage of initially low achieving pupils are rearranged in a higher position, and thus many among them are characterised as average according to contextual models.

3. 37. Are classrooms consistently effective according to a range of models?

This section will shed light on the extent to which classrooms perform consistently across different models, or according to different performance criteria. The following analysis will reveal the extent to which classrooms perform consistently when unadjusted classroom effects, adjusted (or net) classroom effects or contextual effects are examined. This analysis will be pursued for both mathematics and language and it will examine how stable are the classroom rankings, which are derived from different models. For example, do classrooms scoring highly in terms of pupils' final attainment scores also score highly in terms of progress? The computation of the correlation coefficients between classroom residuals derived from a set of models at a time (for example according to the unadjusted absolute attainment model and according to the unadjusted progress model) will shed some light on this issue. Furthermore, these correlation coefficients will be compared between themselves (with correlation coefficients derived from other pairs of models).

In the next section, correlations between the classroom rankings derived from a range of models a) for mathematics and b) for language are shown through estimating correlation coefficients between pairs of models.

For the calculation of classroom residuals, the following models were employed:

- a) A Null model.
- b) An absolute attainment model accounting for pupils' characteristics.
- c) An unadjusted progress model.
- d) A progress model additionally adjusting for pupils' characteristics, and
- e) A contextual model.

These correlation coefficients show that even when the employment of a pair of criteria such as 'unadjusted attainment' and 'attainment adjusted for social class and ethnicity' shows that classroom rankings are proximal, it is probable that classroom rankings are disparate when other criteria are employed. Even if certain classroom rankings are proximal according to the null model and an absolute attainment model (for example a model additionally adjusting for pupils' social class and ethnic background), it is likely that these classroom rankings will not remain proximal when they are derived from a progress model or from an adjusted progress model, and vice versa.

Thomas (1995) suggests that the more detailed information available, the lower the correlation between 'raw' and 'adjusted' results.

Table 3. 17: Correlations of intercept residuals derived from the null, the progress, and the progress model further adjusting for individual pupils' characteristics and the contextual model for mathematics.

Model – types:	Null model 2	Absolute attainment model 8e (Adjusting for ethnic groups & gender)	Progress model 2	Progress model 3d2 ²⁴	Contextual model 3_cont_1 ²⁵
Null model 2					
Absolute-attainment model 8e (Adjusting for ethnic groups & gender)	0.98				
Progress model 2	0.63	0.64			
Progress model 3d2 adjusting for initial attainment score and individual pupils' characteristics	0.62	0.63	0.98		
Contextual model 3_cont_1	0.53	0.54	0.96	0.97	

The above five models can be considered to represent different model-types.

24 Progress model 3d2 controls for pupils' initial attainment score and individual pupils' characteristics (social class, ethnicity and the interaction term 'being foreign/repatriated by initial attainment score').

25 Contextual model 3_cont_1 controls for initial attainment score, ethnicity and the 'percentage of pupils in the class who are foreign/repatriated and low achievers'.

As we carry on model building classroom residuals derived through successive modelling stages (e.g. residuals derived from a contextual model and residuals derived from a progress model that further adjusts for individual pupils' characteristics, or residuals derived from the unadjusted progress model and residuals derived from a progress model that further adjusts for individual pupils' characteristics) approximate each other. This is shown in the increase in the correlation coefficients between classroom residuals derived through successive stages of modelling. The correlation coefficients in the above table show that there is more proximity between classroom rankings derived from null model 2 and adjusted absolute attainment model 8e controlling for individual pupils' characteristics ($r = 0.98$) or between an unadjusted progress model and progress models additionally controlling for individual pupils' characteristics ($r = 0.98$) or between an unadjusted progress model and contextual models ($r = 0.966$), than between an unadjusted progress model and the null model ($r=0.63$) or between the unadjusted progress model and an absolute-attainment model additionally adjusting for individual pupils' characteristics ($r=0.63$).

Hence classroom-intercept residuals derived from a contextual model are very close to classroom-intercept residuals derived from progress model 2 or from a progress model adjusting for individual pupils' characteristics (3d2). On the other hand, intercept residuals derived from progress model 2 are less close to intercept residuals derived from null model 2. Intercept residuals derived from the contextual model (3_cont_1) are less close to intercept residuals derived from null model 2 or to intercept residuals derived from an absolute-attainment model controlling for individual pupils' characteristics.

Low correlation coefficients between classroom residuals derived from two model types imply that classroom rankings according to the two criteria are dissimilar, whereas high correlation coefficients show that classroom residuals derived from these models are similar.

Table 3. 18: Correlation coefficients of intercept residuals derived from the null, the progress, and the progress model adjusting for individual pupils' characteristics and the contextual model for language.

Model – types:	Null model 2	Absolute attainment model 3b2 (Adjusting for gender, ethnicity and social class)	Progress model 4	Progress model 4b ²⁶	Contextual model cont3 ²⁷
Null model 2					
Absolute attainment model 3b2 (Adjusting for gender, ethnicity and social class).	0.887				
Progress model 4 (controlling for initial attainment score)	0.410	0.515			
Progress model 4b with individual pupils' characteristics.	0.416	0.531	0.997		
Contextual model cont3.	0.196	0.341	0.962	0.958	

The above five models can be considered to represent different model types.

26 Progress model 4b controls for initial attainment score and individual pupils' characteristics (gender and for pupils from the former USSR)

27 Contextual model cont3 controls for pupils' initial attainment score, individual pupils' characteristics (social class, the interaction term 'social class by initial score') and contextual classroom characteristics ('percentage of low achieving pupils in the class').

For language as well, the proximity of classroom residuals derived from a range of models was examined through the examination of correlation coefficients between a set of models at one time. Similar conclusions to those derived from the examination of classroom residuals for mathematics can be drawn.

One reason why some of these correlations are high is the following: additional explanatory variables measuring individual pupils' characteristics do not vary greatly between classrooms.

Two models are being built, with and without the explanatory variable x_{2ij} :

$$y_{ij} = \beta_0 + \beta_1 x_{1ij} + \beta_2 x_{2ij} + u_{1j} + e_{ij}$$

and

$$y_{ij} = \beta_0^* + \beta_1^* x_{1ij} + u_j^* + e_{ij}^*$$

Suppose

$$x_{2ij} = x_2 + u_{2j} + e_{2ij}$$

and that

$$u_j^* \approx u_{1j} + \beta_2 u_{2j}$$

$$u_{1j}, u_{2j}$$

if

$$\text{cov}(u_{1j}, u_j^*) = \sigma_{u1}^2$$

Then according to Goldstein (*personal communication*) the correlation between these two sets of intercept residuals becomes:

$$\rho(u_{1j}, u_i^*) = \frac{\sigma_{u1}^2}{\sigma_{u1}(\sigma_{u1}^2 + \beta_2^2 \sigma_{u2}^2)^{1/2}}$$

If the between-classrooms variation of the extra variable x_{2ij} is $\sigma_{u2}^2 = 0$, then the above correlation between the intercept residuals derived from the two models becomes 1.

If this between-classrooms variation for the second variable is other than zero, then the denominator is greater than the numerator and the correlation decreases as σ_{u2}^2 increases.

The bigger the between-classrooms variation for the second variable, the smaller the correlation coefficient between the sets of residuals derived from the two models.

3. 38. Consistency between outcomes

To address the issue of whether classrooms perform consistently across two different domains, whether classrooms' rank order for mathematics is congruent with their respective rank order for language is examined (Scheerens and Bosker, 1997).

Classroom effects for mathematics and for language were contrasted according to four models: unadjusted attainment (null) models, unadjusted progress models, progress models adjusted for individual pupils' characteristics and contextual models.

Table 3. 19: Correlation coefficients between mathematics and language intercept residuals derived from null, progress, progress models adjusted for individual pupils' characteristics and contextual models.

	Model – type:	Language			
		Null model 2	Progress model 4	Progress model 9a ²⁸	Contextual model cont3 ²⁹
Mathematics	Null model 2	0.652** p. <0.001			
	Progress model 2		0.167 p. = 0.14		
	Progress model 3d2 with individual pupils' characteristics (Controlling for social class, ethnicity and the interaction term 'foreign by initial score')			0.195 p. = 0.09	
	Contextual model 3_cont_1 (Controlling for pupils' initial attainment score, minority status, the interaction term 'minority status by initial score' and the 'percentage of pupils in the class who are foreign and low achievers')				0.180 p. = 0.12

The above table shows that the only significant correlation coefficient between classroom-intercept residuals derived from similarly specified models for mathematics and for language is that between null model 2 for mathematics and null model 2 for language ($r = 0.652$). This can be characterised as moderately high. Classrooms tend to perform similarly in terms of pupils' unadjusted final performance for both subjects.

28 Progress model 9a controls for initial attainment score, pupils from the former USSR, gender, social class and the interaction term 'social class by initial score').

29 Contextual model cont3 controls for former USSR, gender, social class, social class by initial score and the percentage of low achievers in the class).

The remaining correlation coefficients between pairs of intercept residuals derived through different models for mathematics and for language did not reach significance level (at $p < 0.05$). Hence, when other criteria were employed (progress, progress adjusted for individual pupils' characteristics, or progress adjusted for contextual classroom characteristics) classrooms did not perform similarly across subjects.

However the above correlations are biased since they are estimated using the estimated residuals and hence the significance tests are not very meaningful. A full multivariate analysis would provide unbiased estimates of these correlations, but given the limitations of this study such an analysis was not undertaken.

Teddlie et al. (2000, p. 117) reported that '*evidence from the US, the UK and the Netherlands indicates a moderate level of consistency of school effects across measures of achievement from different subject areas, which typically involve assessment in basic skills (e.g. mathematics, reading, writing)*'. Also Teddlie et al. (*op. cit.*), Mandeville and Anderson (1987) and Mandeville (1988) reported moderately strong, positive correlations between correlation coefficients derived from the same grade for mathematics and reading in elementary schools in South Carolina. These correlation coefficients were in the (0.60 to 0.70) range. These correlations are very much in line with the correlation between classroom intercept residuals derived from null models for language and mathematics ($r = 0.652$) reported in the current study. In addition, research (Bosker, 1989 in Teddlie et al., *op. cit.*) conducted in Dutch primary schools indicated also moderately strong positive correlations ($r=0.72$).

In Britain, Mortimore et al. (1989, p. 204) showed that *'on the whole, schools, which promoted good progress for mathematics, tended also to do so in reading. There were weaker relationships between school effects on progress in writing and progress in other cognitive areas'*. Mortimore et al. (1988) reported positive correlations between school effects on mathematics and writing ($r=0.28$) and mathematics and reading ($r=0.41$). Sammons et al. (1993) reported a correlation of 0.61 between schools' value-added outcomes in reading and for mathematics in year 5, while Strand (1997) reported a correlation of 0.64 between schools' value-added outcomes in reading and for mathematics in national Key Stage 1 tests.

Based on evidence derived from the UK indicating that correlations between school residuals for different subjects are moderately correlated, Fitz-Gibbon (1995a, p. 3, cited in Teddlie et al., *op. cit.*) suggests that data (related to school or to classroom effects) need to be published separately, by subject, instead of test scores being aggregated across subjects.

If not all the pupils who sat the mathematics exam sat the language exam at either of the testing occasions (at the beginning or at the end of the school year), more accurate estimates of correlation coefficients between classroom intercept residuals would have been obtained if a bivariate analysis had been undertaken (Goldstein, 2003). Bivariate analysis allows pupils' outcomes to be modelled jointly in two or more subjects.

Bivariate analysis is valuable when there are many missing cases in the dataset, as frequently occurs in longitudinal studies. If a bivariate analysis had been employed, correlation coefficients between intercept and slope residuals derived from mathematics and from language would have been estimated with greater precision. A bivariate analysis

was not considered to be necessary here, as there were few missing test scores for pupils in the current dataset. As pupils' testing on the performance-based tests both at the beginning and at the end of the school year was carried out during the same day in classrooms, the vast majority of pupils who sat the language test also sat the mathematics test. Therefore, it is reasonable to calculate Pearson's correlation coefficients between classroom intercept residuals derived from similarly specified models for mathematics and language.

3. 39. *School Effectiveness Research and concerns of Accountability*

'Accountability pressures have forced most education systems to press ahead with large-scale assessment programmes' (Rowe, 2000, p. 75). All educational systems have vested interests in producing acceptable standards for all students enrolled, thus in the-ory not allowing attainment levels among schools to fall below a certain level. The fa-voured measures to be used to evaluate the performance of classrooms or schools are performance indicators (PIs). Classroom or school rankings in terms of attainment, at-tainment adjusted for conditions of disadvantage, progress, and progress adjusted for conditions of disadvantage can be perceived as a set of performance indicators, inform-ing policy makers and the schools about the status, potential problems, and relative strengths and weaknesses of classrooms or schools. *'A PI can be defined as an item of information collected at regular intervals to track the performance of a system'* (Fitz-Gibbon, 1990, p. 1, cited in Sammons, 1994). *'The collection and use of performance indicator information is intended to promote accountability, and to help identify areas where improvement is required at a variety of levels: district, individual school or classroom. Performance indicators need to be collected on a regular basis to allow the monitoring of changes over time'* (Sammons, 1994, p. 33).

In England, students' outcomes are regularly monitored, as pupils are yearly assessed in basic skills at the end of Key Stage 1, 2, 3, and 4 (at the end of 2nd, 6th, 10th and 12th grades, and at ages 7, 11, 14 and 16). *Average school performances* are published an-nually in the so-called 'league tables'³⁰. *'The systematic publication of 'performance*

³⁰ Average school performance is published on a school by school basis by the Department for Educa-tion and Skills, and this information is ranked-ordered into League Tables.

tables' for Key Stages and exam results began in 1992 and is now an established feature of the educational system in England and Wales' (Goldstein, 2001). This exercise aiming to monitor more closely schools' educational outcomes is of increased importance for schools in England as on many issues (such as on equal educational opportunities) schools define their own policies and emphasise different priorities for action, demonstrated in schools' policy documents. Hence, the British state strives to ensure that all schools are able to provide a basic standard in education. However, Goldstein (2000, p. 1) strongly criticised the publication of league tables, on the premise that 'League tables are a poor method of ensuring accountability, can distort teaching and are a poor way to measure 'standards'.

Progress criteria reflect more school or classroom endeavour to raise pupils' attainment. Gray, et al., (1999, p.168) described effectiveness as the extent to which the school boosted pupils' final score performances above the levels that they should have predicted from knowledge of their starting points. Willms (1992, p. 34) stated that: *'A preferable indicator of a school's performance is the distribution of the rates of growth of its pupils, rather than the distribution of pupils' scores on one occasion'.*

L. Stoll and P. Mortimore, (1997, p. 9) defined an effective school as *'one in which pupils progress further than might be expected from consideration of its intake'* (in comparison with other schools serving pupils with similar baseline attainment and other socio-economic characteristics).

SER has strongly criticised unadjusted comparisons between schools as published in league tables in England, as the positions of schools according to league tables reflect schools' relative rankings based on absolute attainment measures of pupils' test scores. *'The apparent simplicity of rankings of average student test and exam results is deceptive: they largely reflect 'intake' achievements and, at the very least, we should adjust for intake differences - a value added approach'* (Goldstein, and Spiegelhalter, 1996, p. 1). Unadjusted comparisons between schools are not justified *for accountability purposes*. According to Goldstein (1997, p. 18) *'the principal argument against examination of league tables is that the performance of a school is determined largely by the pre-existing achievements of the students when they enter it'*. A more comprehensive picture of the work undertaken by the school would be in terms of progress that pupils enrolled in the school make during a given period (often one school year). Progress criteria are preferable if the goal is to identify schools or classrooms having performed at an above-average or at a below-average level, as progress criteria reflect more school or classroom endeavour to raise pupils' attainment. Progress comparisons are required in any exercise involving teachers' assessment, as teachers serving in schools and classrooms that in spite of their circumstances managed to contribute to increased progress rates for their pupils should be rewarded, irrespective of pupils' final attainment scores, while other schools/classrooms with below average progress rates should be more carefully monitored. Inhibiting conditions may have prevented schools/classrooms with below average progress rates to provide enough challenge or adequate support to their pupils so as to enable them to raise their final attainment score to a higher level than that indicated by their initial attainment. Hence these schools/classrooms failed to show

adequate rates of progress in their pupils' educational outcomes during the period investigated.

Value-added estimates assess schools' or classrooms' impact in promoting pupils' outcomes because the main contribution of a school or classroom is measured by the degree to which pupils' outcomes are raised during pupils' enrolment in this school/classroom, a fact that especially applies in the case of schools/classrooms with an initially low achieving pupil intake.

According to Myers & Goldstein (1997), even in cases where a value-added approach is adopted, there are problems associated with the presentation of a single figure representing the value-added ranking for each school/classroom, as the school/classroom may show different effectiveness patterns for different pupil groups; in other words, it may be 'differentially effective'. A second caveat is that a large margin of error or uncertainty is attached to school/classroom estimates. A third caveat is that comparisons may be invalid in cases where during the period between the two measurements, a large number of pupils enrolled in the schools examined moved between schools.

An example of sensible use of SER information in the UK is the Hampshire project, Goldstein (2000, p. 5). In this project schools' outcomes in terms of adjusted progress, which are reflected in the rankings of the schools' residuals, have been considered '*as screening devices that provide information alongside other information available that might be of constructive use*' (Goldstein, *op. cit.*).

According to Goldstein et al. (1993), comparisons between individual schools are not acceptable from a statistical point of view. Instead, comparisons of individual school or

classroom outcomes with the average are legitimate. Furthermore, the publications of schools' results in the form of league tables in Britain was based on pupils' unadjusted attainment scores on public examinations at key transition points of pupils' schooling. Since 2002 however, the Department for Education and Science (DfES, 2002) has piloted the publication of information about individual schools based on value – added data, estimating progress made by pupils in each individual school during a given school year. Yet, when accountability is the major concern, adjustments should be made for more individual characteristics such as social class and/or minority status so that schools serving similar pupil intakes are compared (so that like are compared with like). In many cases, it might not be viable to adjust for all correlates that may be perceived as indicators of disadvantage, as there is no simple answer to the question 'which individual pupil or contextual classroom characteristics matter?'

Goldstein (1994, p. 157) comments: *'The unease that many feel about publicly accountable systems seems to stem partly from the current educational climate, where the political intention would appear to be that of punishing the relatively poor performers rather than helping them to understand their situation. Nevertheless, this should not deter us from considering how to set up accountability systems, using the best performance techniques and making all the necessary caveats'.*

The situation is acute in 'failing schools' (schools scoring at a below-average level), a term used by politicians in the UK in order 'to shift responsibility' (Myers and Goldstein, 1997). *'The language and labelling of 'failing schools' adopted in certain countries often exacerbates and prolongs the problem of schools in difficulty'* (Hargreaves, 1997; Myers & Goldstein, 1997 in Stoll, 1999, p. 514).

In England, the Office for Standards in Education, in an endeavour to ensure that all schools provide an adequate standard of education, reinforces '*special measures*' in '*failing schools*' defining priorities for action in these schools so as to improve pupils' performance (OFSTED, 1999). These special measures may be assigned to the school after inspection, and they can vary according to the school's case. However, one of the inspection criteria is how the school performed in annual national assessment exercises and what the school's position is in league tables. Pupils' outcomes are one of the core criteria in OFSTED's (2003) school evaluation schedule along with a cluster of processes related to teaching and learning. Special measures are taken in poorly performing schools, which may among others consist of allocating a grant to the school, '*designed to support activity designed to improve standards of pupil performance in order to meet school, district and national Targets. At least 90% of the money must be devolved to schools to enable them to address priorities identified in their school development plans and post-OFSTED inspection plans*' (DfES, the standards site, 2002). Eventually, if schools performing at a below-average level do not improve their performance over consecutive years, they will have to shut down; in due time, most of them reopen with a new principal, teaching staff and administration.

The regular monitoring of pupils' progress is considered as an effectiveness-enhancing factor by Scheerens and Bosker (1997).

The '*No child left behind act*' (Boehner, United States of America - Congress House, 2001) mentions that districts and schools should monitor the performance of low achieving groups of pupils who suffer from educational disadvantage and take measures to ensure that these groups make adequate progress. Schools can improve the

attainment of educationally disadvantaged pupil groups (those who speak the language of the school as a second language, or those who live in poverty), which can be raised after their needs have been identified. Educational monitoring using progress or adjusted progress indicators enables policy makers to assess these needs in order to intervene where possible and allocate extra teachers and funds to schools serving disadvantaged populations.

Educational monitoring allows districts, principals and teachers to identify areas where classrooms or schools show low progress rates and diagnose early gaps in pupils' performance. These low performing classrooms and schools take part in a school/ classroom improvement programme; underachieving pupils can receive support teaching, or intervention in a support class, teachers can receive in-service training especially in subjects where they demonstrated low results; more broadly, an investigation of the reasons why this low performance occurred takes place, so that subsequent corrective action is undertaken. Monitoring appears as a characteristic of unusually effective schools in Levine and Lezotte (1990) and in Sammons et al. (1995a). According to Cotton (1995), districts, schools and classrooms profit if they make good use of the results of educational monitoring. Educational monitoring based on assessment results can guarantee that the number of pupils who escape district and school safety net is very limited. The ultimate goal is that high levels of attainment in the core subjects be reached, and thus in the long run, educational standards be raised.

School Effectiveness research can be conceived as a methodological tool allowing schools and districts to evaluate themselves and further to identify subject areas in

which the pupils have accumulated weaknesses and / or schools and classrooms with significantly lower progress rates than the average school / classroom. Such an exercise can further identify cases of schools and / or classrooms where there are groups of pupils with significantly lower attainment or progress than the majority group. Such differences can be perceived as systemic weaknesses, the identification of which allows districts to assume a coordinative role in setting up interventions aiming to raise the educational outcomes of schools, classrooms and pupil groups in their jurisdiction.

4. FINDINGS

To start with the current section discusses the finding of the existence of classroom effects and the absence of school effects and their implications for the Greek setting; then the focus of the findings shifts to explaining the impact of individual pupils' characteristics to pupils' final attainment score and pupils' progress rates, when these individual characteristics have been introduced one at a time or jointly in absolute attainment or in progress models; subsequently the impact of the existence of random coefficients associated with these individual pupils' characteristics in attainment or in progress models is discussed; then the impact of significant contextual effects and their implications for classrooms where there is a big concentration of disadvantaged pupils are discussed; finally the above findings are linked with the issue of the limited implementation of the two support schemes which operate in the Greek setting.

4. 1. The existence of classroom effects

According to Willms (1995), classroom effects may be defined as the extent to which attending a given classroom modifies pupils' outcomes. On the other hand, school effects can be defined as the extent to which attending a particular school modifies pupils' outcomes. Evidence from the models described in section 3.21 in the previous chapter suggests that in Greek Primary Education classroom effects can be identified and school effects cannot be identified.

For both subjects school effects were tested according to a variety of models in which the total variation has been divided into three levels (at the school, at the classroom and at the pupil level). Null models, unadjusted progress models, and adjusted progress models, which additionally controlled for additional individual pupils' characteristics, were investigated. None of the above models was an improvement on their respective comparison models at $p < 0.05$. Hence, *no school effects were identified in the current school dataset, which represent state primary schools in the city of Piraeus.*

4. 2. Discussion on the absence of school effects for the current Greek dataset

The existence of classroom effects and the absence of school effects for the Greek dataset can be juxtaposed with the findings derived from the American and English school effectiveness research literature, where school effects characterise these school systems. The current findings are in line with the findings from other SER studies indicating that classroom effects are greater than school effects in explaining students' progress (Creemers, 1994; Reynolds et al., 1996). Scheerens and Bosker (1997, p. 84) suggested that teacher effectiveness might have caused the school differences identified in many SER studies. This might be especially the case in studies where the classroom level is not included in the modelling process, as has happened frequently in the analysis of SER studies in the past. However, even where classroom effects are found, classroom effects should not be interpreted as pure teacher-effects, as teacher contribution is enhanced by school initiatives, such as the provision of support teaching, the deployment of resources at the school's disposal, and school climate.

Teddlie et al. (2000, p. 96) found that units located at a higher level are less variable than units at a lower level (as the former aggregate variation from units at a lower level).

'There is more variance between students than there is between classrooms; there is more variance between classrooms than between schools; there is more variance between schools than between school districts' (e.g. Hill and Rowe, 1996; Tymms, 1993 cited in Teddlie et al., 2000, p. 96).

However, in UK schools, between 8% and 12% of the total variance can be attributed to the schools. This difference between the Greek and British school effectiveness results may be partially attributed to the fact that there are underlying differences between the Greek and British educational systems. Reynolds (1991) predicted that a variety of factors within the British educational system, including the decentralisation of power within the educational system down to the individual level of the school was likely to lead to *'a substantial increase in the variation in quality between individual schools...'* (Reynolds, 1991, p. 5). In addition, the educational system of England and Wales is characterised as a 'decentralised system' by Altrichter and Salzgeber (2000, p. 99). According to a school effectiveness study conducted in urban and rural primary schools in Cyprus (Kyriakides *et al.*, 2000, p. 515) where, similar to the Greek educational system *'the curriculum and time allocation are standardised by the ministry of Education and vary very little'*, according to the null model, 8.5% of the total variance is located at the school level and 13.8 % is at the classroom level. Similarly to the results derived from current study, in the Cypriot system the classroom level is more im-

portant than the school level; yet, in the Cypriot case schools have some impact on pupils' educational attainment.

In the current study the short time span that elapsed between the two testing occasions (six months) for a particular school may be one possible reason why school effects were not identified. Although the study was conducted over one school year (nine months), for any given school the period that elapsed between the two testing occasions was six months. For Teddlie et al. (2000, p. 100), '*longitudinal studies conducted over several years yield larger estimates of the magnitude of school effects, whereas for longitudinal studies operating over a limited time-scale...there is the risk that school effects will be unobserved, underestimated*'. Strand (1997) warned that caution should be exercised in looking at the results from any one year, especially when the sample size is relatively small and because schools can change in terms of their effectiveness (Gray et al., 1996b). Hence, frequent replications of School Effectiveness studies are recommended.

The analysis of the IEA's Second International Mathematics Study (Scheerens et al., 1989) that employed both a cross-sectional and a longitudinal research design showed that in some countries (as for example in Sweden, Finland and New Zealand), '*the simultaneous estimation of school and class effects shows that school effects are virtually non-existent when the class effect has been accounted for*' (Scheerens et al., *op. cit.*, p. 794). According to Reynolds (2000, p. 249) for the case of Sweden, the absence of school effect can be attributed among other things to the historically strong control of the Swedish system over the organisation of schooling and possibly to homogeneous school intakes from administratively defined balanced catchment areas. Reynolds' conclusions were based on the reanalysis of the IEA's Second International

clusions were based on the reanalysis of the IEA's Second International Mathematics Study (Scheerens et al., 1989).

According to Teddlie et al. (2000, p. 115), *'the country in which SER occurs certainly has an effect upon the magnitude of school effects'*. Reynolds et al. (2002, p. 277) suggest that *'there are interesting variations between countries in the reliability of their educational systems, with some evidencing 'low variance' and some, predominantly Anglo-Saxon societies, showing larger between classroom and school variability'*.

Macbeath & Mortimore (2001, p. 14) reported the results of Scheerens et al. (1989) from the second International Mathematics and Science study. Scheerens found that in countries with vertically differentiated educational systems such as Belgium and the Netherlands, that is systems in which pupils entered different tracks or schools at a given age, there were large differences in the mean achievements of pupils across schools. In contrast, in more horizontally integrated, or 'comprehensive' systems where pupils moved up together within the same structure (Scottish, American, Swedish, Finnish and New Zealand schools, for example), there were relatively small differences between schools but relatively big differences between classrooms within schools. France, Canada and Israel belong to this second group insofar as there is relatively little variance between schools, but they also reveal comparatively little variance among classrooms within schools.

Another possible reason that may account for the absence of school effects in Greek state schools is that schools do not apply any selective policy in their enrolments. Pupils usually enrol in the state school located closest to their current residence. Experimental

schools, which enrol pupils from the wider geographical area, draw lots among the applicants' names to decide on their pupil intake.

One can attribute the absence of school effects in this Greek study to the fact that districts alone appoint teaching staff to state schools. Teachers can only apply for geographical areas while they are not allowed to apply for particular schools they would prefer to teach in. On the other hand principals and teachers already appointed in state schools do not have any right to select and appoint teachers in their schools. Thus the classroom effect (the existence of significant classroom variation) identified in the current study can be attributed to the fact that teachers are appointed in primary state schools in a random way, which means that very skilled and less skilled teachers may serve in the same school. *'The Greek educational system like many other educational systems in the world is operated by a wide-ranging bureaucracy'* (King, 1983). According to Iliou (1982, 1988) *'the centralised state system of education has guaranteed equality of opportunity to all children, but this has become equated with centralisation, repetition and uniformity in education'*. On the one hand, the fact that school effects were absent among Greek state schools in Piraeus may reflect a less diversified and more egalitarian educational system with better educational opportunities for the majority of schools' population in relation to other countries in which school effects are found. However, a more egalitarian educational system may also entail limited classroom effects, as classroom effects suggest that there are classrooms performing at a below-average level, thus creating less progress for pupils enrolled in them than the majority of classrooms. The existence of classroom effects and the fact that classrooms are

differentially effective for high and low achievers indicates that different classroom settings provide different educational opportunities to different pupils' groups.

Couloubaritsis and Kavouri (1996), in an attempt to study the teaching practices of Greek primary school teachers, found that teachers attempt to ensure uniformity within their schools and classrooms for two reasons. Firstly they have come to regard the Greek Ministry of Education as the only body which tells them what they must and must not do, and secondly, teachers use routines as this guarantees a control over knowledge and therefore over their classroom. However, the existence of classroom effects shows that there are substantial differences in progress between classrooms, which suggests that teachers' contributions in raising pupils' outcomes vary.

Gotovos (1986) argued that *'a routine logic prevails in the bureaucratic Greek school system. Schools tend to be static, each with the same conditions and range of activities, the same prospect of everyday events, formal teaching with the same routine actions, a lack of communication between staff members, principal and pupils, absence of extra-curricular activities as well as teachers continuing with their own traditional practices'*. However, this static picture has started to change nowadays with the initiation and institutionalisation of support schemes catering for the needs of foreign/repatriated pupils such as the reception classes and coach classes (journal of the Greek Government, law 1894/ ΦΕΚ 110/27-8-1990) as well as with the institutionalisation of the whole day school or support with homework sessions (law 2525/97 in Journal of the Greek Government, ΦΕΚ 188/2Α' article 4). Through the adoption of these initiatives there is new room for Greek state schools to implement their own policies on support

4. 3. Estimating the impact of individual pupils' characteristics on pupils' final attainment scores and progress rates

In the next stage of the analysis models adjusting for individual pupil characteristics are tested. While 'absolute attainment models' adjust only for individual pupils' characteristics but not for pupils' initial attainment score, progress models adjust for both pupils' initial attainment scores and for other individual pupils' characteristics. '*Absolute attainment models*' assess the impact of individual pupils' background factors on their final attainment score. These models address the following research questions: Are there any attainment differences between pupil groups defined by ethnicity, social class and gender? Is there any evidence of lower attainment for disadvantaged groups of pupils? To what extent do additional explanatory variables (such as the educational level reached by a pupil's mother, hours of weekly homework undertaken by a pupil, days that each pupil has been absent during the school year, the number of friends a pupil has within the class and whether a pupil has attended nursery school or not) have an impact on pupils' final attainment scores? By controlling for pupils' initial attainment scores, the effect of individual pupils' characteristics on pupils' progress during the period between the two measurements is estimated. These models address the following research questions: Are there any differences in progress between pupil groups defined by ethnicity, social class and gender? To what extent do the previously mentioned additional explanatory variables have an impact on pupils' progress made during the school year? 'Adjusted attainment models' and 'adjusted progress models' that follow control for a set of individual pupils' characteristics so as to perform adjusted comparisons in terms

of pupils' final attainment scores. For example, when an 'adjusted attainment' model controls for ethnicity and social class, classroom heterogeneity attributed to ethnicity and social class is removed (from the level 2 random part) and thus classroom residuals derived from such a model enable comparisons of attainment between similar classrooms in terms of pupils' ethnicity and social class to be undertaken. When an 'adjusted progress' model controls for initial attainment score, ethnicity and social class, classroom heterogeneity attributed to ethnicity and social class is removed (from the level 2 random part) and thus classroom residuals derived from such a model enable comparisons of progress between similar classrooms in terms of pupils' ethnicity and social class to be undertaken. At the same time the coefficients of social class and ethnicity in the fixed part of the models show the average effect of these two variables on pupils' final attainment scores or on pupils' progress, compared to the base category, having controlled for the effect of the second variable.

The joint impact of individual pupils' characteristics on pupils' progress was examined by including combinations of background factors in a model that already adjusts for pupils' initial attainment.

It is generally recognised in studies adhering to the SER paradigm that *'pupils' prior attainment has by far the largest impact on their later attainment (particularly when a finely graded measure is employed)'* (Thomas and Mortimore, 1996: 6), more than any other independent combination of variables measuring individual pupils' characteristics or contextual characteristics.

According to '*progress model 2*', for mathematics and '*progress model 4*' for language, shown in sub-sections 3.10 and 3.14 of the current thesis, pupils whose initial attainment score differs by one normal score unit are expected to differ by 0.816 normal score units in mathematics and by 0.92 in *language*. Pupils' initial attainment scores have a differential impact across classrooms, as initial attainment score has a random coefficient assigned to it at the classroom level, for both subjects according to '*progress model 2*' for mathematics and '*progress model 4*' for language. Classrooms differ in the amount of progress they contributed for high and low achievers.

The joint impact of individual pupils' characteristics on pupils' final attainment scores and on pupils' progress rates was examined by including combinations of background factors in an 'absolute attainment' or in a progress model. Strand (1999) suggested that in the modelling process combinations of background factors should be considered defining pupil groups whose educational progress or attainment is of particular concern.

Often the coefficient of a variable that has been introduced initially in the modelling process in an absolute attainment or in a progress model diminishes when the model adjusts for an additional variable. Then the analyst can compare the coefficient of the first variable in the model where the additional variable has not yet been introduced in an attainment or in a progress model, with the coefficient of this variable in the model where the additional explanatory variable has been introduced. By introducing one additional explanatory variable at the time in a model that already adjusts for the effect of the first variable, one can identify factors that can alleviate the impact of the first variable on pupils' attainment or on pupils' progress.

Hair et al. (1995, p. 124) writing about regression analysis suggest that analysts must evaluate not only one final regression model but also the potential independent variables that were omitted if a sequential search or combinatorial approach was employed, as multicollinearity may substantially affect the variables ultimately included in the regression.

In the following tables 4.1 to 4.4 the main effects of individual pupils' characteristics on pupils' final attainment scores and on pupils' progress and the percentage of reduction of these main effects in the presence of an additional variable are shown.

Hair et al. (1995, p. 124) suggest that the standardized coefficients derived from a regression model should be used as a guide to the relative importance of individual independent variables only when collinearity is minimal. The values of the standardized coefficients can be interpreted only in the context of other variables in the equation. In the current study the values of the coefficients derived from adjusted attainment models are significant at $p < 0.01$. The values of the coefficients derived from adjusted progress models are significant at $p < 0.05$, unless another p-value is provided. Standard errors are shown in brackets, under the main effects.

Table 4. 1. Adjusted attainment models: The main effects of variables measuring individual pupils' characteristics (in the diagonal); and the percentage reduction of the main effects (in the rows, off the diagonal) when the model controls for a second explanatory variable³¹

a) Mathematics

	Minority Status	Friends	Social class	Gender	Nursery	Homework
Minority Status (1: foreign, 0: Greek)	-0.468 (0.070)	49.5%	38.4%	5.5%	26.9	16.4%
Friends	4.1%	0.148 (0.008)	10.1%	-1.3%	2.7%	8.8%
Social class (1: lower, 0: higher)	4.4%	36.9%	-0.412 (0.052)	0.4%	6%	12.1%
Gender (1: boy, 0: girl)	10.6%	-14.3%	6.5%	-0.245 (0.048)	16.3%	32.6%
Nursery school attendance (1: attended Nursery, 0: did not attend)	20.3%	37.4%	6.7%	5%	0.379 (0.069)	8.9%
Homework Hours of study per week	0%	16.6%	4.2%	0%	0%	0.024 (0.003)

Table 4. 2. Adjusted attainment models: The main effects of variables measuring individual pupils' characteristics (in the diagonal); and the percentage of reduction of the main effects (in the rows, off the diagonal) when the model controls for a second explanatory variable

b) Language

Models:	Minority Status	Friends	Social class	Gender	Nursery	Homework
Minority Status (1: foreign, 0: Greek)	-0.873 (0.068)	22.6%	13.8%	7.6%	18%	8%
Friends	12.2%	0.139 (0.008)	11.5%	-3%	6.4%	9.3%
Social class (1: lower, 0: higher)	6.8%	32.5%	-0.510 (0.060)	-1.3%	8.2%	13.3%
Gender (1: boy, 0: girl)	8.3%	-5.3%	1.5%	-0.579 (0.047)	6.9%	11%
Nursery school attendance (1: attended Nursery, 0: did not attend)	27.4%	20.3%	15.7%	8.1%	0.604 (0.068)	7.8%
Homework (hours of study per week)	4.5%	13.6%	13.6%	9%	4.5%	0.022 (0.003)

³¹ The percentage of reduction is calculated from the absolute attainment model as the initial value of the coefficient of interest minus the final value of the coefficient of interest when the model additionally adjusts for a second variable, divided by the initial value of the coefficient of interest. This method is portrayed by Ian Schagen (personal communication) and is going to appear in his forthcoming book.

Table 4. 3. Adjusted progress models: The main effects of variables measuring individual pupils' characteristics (in the diagonal); and the percentage of reduction of the main effects (in the rows, off the diagonal) when the model controls for a second explanatory variable.

a) Mathematics

Models:	Minority Status	Friends	Social class	Gender	Nursery	Homework
Minority Status (1: foreign, 0: Greek)	-0.113 (0.043)	30%	22.1%	4.4%	NS	NS
Friends	3%	0.033 (0.006)	6%	-3%	NS	3%
Social class (1: lower, 0: higher)	10.5%	23.7%	-0.076 (0.032)	NS	NS	NS
Gender (1: boy, 0: girl)	9.2%	29.6%	51.8%	-0.054 (0.028) P=0.061	NS	27.7%
Nursery school attendance (1: attended Nursery, 0: did not attend)	NS	NS	NS	NS	NS	NS
Homework (hours of study per week)	0	0	0	NS	NS	0.003 (0.002)

Table 4. 4. Adjusted progress models: The main effects of variables measuring individual pupils' characteristics (in the diagonal); and the percentage of reduction of the main effects (in the rows, off the diagonal) when the model controls for a second explanatory variable.

b) Language

Models:	Minority Status	Friends	Social class	Gender	Nursery	Homework
Minority Status (1: foreign, 0: Greek)	NS	NS	NS	NS	NS	NS
Friends	NS	0.013 (0.004)	NS	-30.7%	NS	NS
Social class (1: lower, 0: higher)	NS	NS	NS	NS	NS	NS
Gender (1: boy, 0: girl)	NS	-15.6%	NS	-0.096 (0.021)	NS	NS
Nursery school attendance (1: attended Nursery, 0: did not attend)	NS	NS	NS	NS	NS	NS
Homework (hours of study per week)	NS	NS	NS	NS	NS	NS

In the previous tables (4.1. to 4.4.) the diagonal shows the main effects of variables measuring individual pupils' characteristics when these variables are introduced one at a time; in the cells below and above the diagonal is shown the percentage of reduction of the main effects of variables represented in the rows when the model controls for an additional explanatory variable shown in the respective column; For example, when the absolute attainment model adjusts for minority status and the number of friends, the initial coefficient of minority status in the model adjusting for minority status only (-0.468) reduces by 49.5% (table 4.1.).

Sammons (1995, p. 479) suggested that a positive way forward would be to compare the impact of background factors for reading and for mathematics, shown in fixed part coefficients. Sammons identified the point that the impact of background factors (gender and ethnic effects) is stronger for reading than for mathematics attainment at junior school. Therefore the magnitude of main effects can also be contrasted across subjects. For example the impact of minority status on pupils' final attainment scores for mathematics, which is shown in the first diagonal position in table 4. 1. can be compared with the impact of minority status on pupils' final attainment scores for language which is shown in the first diagonal position in table 4.2. In addition, the impact of gender on pupils' progress for mathematics, which is shown in the first diagonal position in table 4. 3. can be compared with the impact of gender on pupils' progress for language which is shown in the first diagonal position in table 4. 4. These comparisons are possible, given that both final attainment scores (for mathematics and language) are expressed in normal score units.

Table 4. 5. Examining the impact of gender, social class and ethnicity on pupils' attainment and progress rates on both subjects.

	Coefficients & Standard errors			
	Mathematics		Language	
Models:	Adjusted attainment model with gender, social class and <i>minority status</i>	Adjusted progress model with initial attainment score, social class and <i>minority status</i>	Adjusted attainment model with gender, social class and <i>minority status</i>	Adjusted progress model with initial attainment score, gender and former-USSR ethnic group
Model number:	6b3	3d	6b3	4b
p-value:	0.0008	0.034	0.001	0.0058
Constant	0.391 0.055	0.048 0.034	0.761 0.059	-0.049 0.023
<i>Initial attainment score</i>		0.802 0.023		0.903 0.015
<i>Gender</i> (1: boy, 0: girl)	-0.207 0.050		-0.528 0.047	-0.094 0.021
<i>Social class</i> (1: lower, 0: higher)	-0.397 0.052	-0.068 0.032	-0.484 0.055	
<i>Minority status</i> (1: foreign, 0: Greek)	-0.261 0.078	-0.088 0.048	-0.684 0.073	
Former-USSR ethnic group (1: pupil from former-USSR, 0: Greek pupil)				-0.165 0.060
Random terms:				
σ_{u0}^2 variance of the intercept at classroom level	0.076 0.020	0.049 0.011	0.048 0.015	0.023 0.005
σ_{u01} Covariance of the intercept with initial score		0.003 0.006		0.009 0.003
σ_{u1}^2 variance of initial score at classroom level		0.018 0.006		0.006 0.003
σ_{e0}^2 variance of the intercept at pupil level	0.817 0.032	0.281 0.012	0.722 0.029	0.140 0.007
σ_{e1}^2 pupils' variance of the slope				0.024 0.004
σ_{e01} Covariance of the intercept with initial score				0.024 0.006
-2loglikelihood	3627.1	2243.6	3438.6	1602.1
N (cases used)	1348	1333	1347	1503
All pupils were:	1858	1858	1858	1858
VPC (%)	8.5	14.8	6.2	14.1

In the previous table (4.5.) the extent to which ethnic group membership, social class and gender account for variations in pupils' attainment and progress rates in language and in mathematics is examined. The above table performs adjusted comparisons in both pupils' final attainment scores and progress rates. The first two columns of table (4.5.) refer to attainment and progress differences between pupil groups for mathematics, whereas the last two columns show attainment and progress differences for language. Overall, nearly 8.4% of the total variation in pupils' mathematics performance and 22.3% of the total variation in pupils' language performance was accounted for by these three individual characteristics variables. There was a slightly greater reduction in classroom level variation for Greek (29.4%) than for mathematics (26.9%) after accounting for these three individual pupils' characteristics, in the absolute attainment models in relation to their respective null models. A plausible reason for this is that pupils are more likely to learn language at home and be influenced by social background factors, while in a classroom initially highly attaining pupils progress more than average and low attaining pupils. These three background factors were better predictors for language than for mathematics final attainment scores at the end of primary school. It is clear that all three variables have a significant impact on pupils' final attainment scores for both mathematics and language. Boys perform less well than girls, pupils from lower social class perform less well than pupils from higher social class and foreign and repatriated pupils perform less well than majority pupils in both mathematics and language. Ethnicity and social class have a joint impact on pupils' progress in mathematics; pupils from lower social class make less progress than pupils from higher social class and pupils from foreign/repatriated ethnic background make less progress than

pupils from Greek background. In addition, progress differences can be jointly identified for both boys and pupils from the former USSR ethnic group for language. Boys and pupils from the former USSR make less progress than girls and Greek pupils respectively. However, the reader should refer to the attainment tables 4. 1. and 4.2 as well as to the progress tables 4.3. and 4.4., in order to investigate the joint impact of combinations of two background characteristics at a time on pupils' final attainment scores and progress rates. More specifically the reader may find out the extent to which the impact of a given variable (the coefficient of which is shown in the diagonal of each table) is reduced when the attainment or the progress model additionally adjusts for a second variable measuring individual pupils' characteristics, and whether combinations of two characteristics produce significant coefficients in absolute attainment or progress models.

4. 3. 1. Differences in attainment and in progress between foreign/repatriated and majority pupils

First of all, the impact of belonging to a different **ethnic group** than the majority group is being assessed based on attainment and progress models where ‘minority status’ is used as a sole explanatory variable. These coefficients are shown in the first square of the diagonal of tables 4.1. – 4.4. native Greek pupils serve as the baseline group. The differences to be mentioned are significant at $p < 0.05$, unless otherwise noted.

In terms of absolute attainment, for both mathematics and language, belonging to a foreign ethnic group has a negative impact on pupils’ final attainment scores:

Foreign and repatriated pupils scored significantly less than the majority group. When introduced as a sole explanatory variable, the difference in attainment was -0.468 normal score units for *mathematics* and -0.873 normal score units for *language*.

While for mathematics foreign/repatriated pupils made *less progress* than Greek pupils they did not seem to make less progress than their Greek counterparts in language. For mathematics foreign/repatriated pupils scored -0.113 normal score units ($p = 0.043$) less than the majority group, as shown in table 4.3.

This demonstrates that when the study was conducted substantial underachievement for foreign/repatriated pupils in relation to majority pupils existed at the end of primary school (6th grade). The effect that belonging to a foreign/repatriated ethnic group has on pupils’ final attainment scores is almost double in magnitude for language than it is for mathematics. Being a foreign/repatriated pupil has a negative impact on pupils’ final attainment scores even when the absolute attainment model adjusts for an additional

variable, measuring any other individual pupil's characteristic. However, adjusting for an additional explanatory variable reduces the coefficient of ethnicity. Haque & Bell (2001) reported that the magnitude of attainment and progress differences between the ethnic groups participating in their study was substantially reduced when a set of different variables such as recency of pupil's arrival, father's occupation, and mother's education were included in an absolute attainment or in a progress model.

In the current study the percentage of reduction of the coefficient of minority status in the presence of an additional variable in absolute attainment terms is shown in Tables 4.1. and 4.2, in the first line above the diagonal. For both mathematics and language absolute attainment models adjusting for ethnicity and an additional variable, are all significant improvements on the absolute attainment model adjusting for minority status only at $p < 0.001$. When the model initially adjusting for minority status additionally adjusts for the number of friends that a pupil has then the impact of being from a foreign/repatriated ethnic background takes its lowest value. For mathematics, its impact is reduced by 49.5% and for language by 22.6%. Friends alleviate the negative effect that being foreign/repatriated has on pupils' final attainment scores. The more friends a foreign pupil has, the more his/her performance is expected to improve, as the use of Greek (the second language) is facilitated.

In terms of progress for *mathematics*, the coefficient of minority status turns out to be non-significant when the progress model additionally controls for the number of friends a pupil has. The coefficient of minority status decreases from -0.113 ($p = 0.009$) in the progress model where it is introduced as a sole explanatory variable, and becomes marginally significant -0.079 ($p < 0.065$), when the model controls for the number of friends

a pupil has. Having friends alleviates the negative impact of ethnicity on the progress rates of foreign/repatriated pupils in relation to the majority group.

A plausible explanation for that may be that the number of friends a pupil has is an integrative factor for minority pupils. An issue that merited investigation is whether foreign/repatriated pupils have fewer friends and therefore the coefficient of minority status diminishes, because there is common variation explained by both the 'minority status' and 'friends' variables. Indeed, an independent samples t-test with '*the number of friends a pupil has*' as the dependent variable and being of foreign/repatriated or of Greek background as grouping variable produced a statistically significant ($p < 0.001$) difference of 1.72 friends between the friends acquired by Greek pupils and foreign/repatriated pupils. On average the number of friends was 2.7 for foreign/repatriated pupils and 4.4 for majority pupils. Maybe foreign pupils have more opportunities to practice and speak the target language if they have native friends. Therefore Swain (1978c, cited in Swain, 1982, p. 51) advocated that '*Sustained interaction of ethnic minority pupils with native peers should be encouraged if the immersion (L2) children are to attain native-like speaking abilities*'.

Tables 4. 1. and 4. 2. show that the negative effect that being of foreign/repatriated ethnic background has on pupils' final attainment scores is reduced as well when Nursery school attendance is included in the same model with ethnicity. Then the magnitude of the minority status coefficient reduces by 26.9% for mathematics and by 18% for language. Having attended Nursery school alleviates the negative effect of being of minority status in both subjects. This is expected as nursery school offers socialising experiences and an additional exposure to a language rich environment, which offers opportu-

nities to experiment with materials and learn the language, which is used as a medium of instruction through pupils' engagement in various manipulative activities.

For mathematics, in terms of absolute attainment (model 17th) *Albanian* pupils and foreign pupils grouped in the '*other*' category performed less well than majority Greek pupils. However, the coefficient of the pupils coming from the democracies of the former Soviet Union is negative, but of borderline significance at ($p=0.061$).

For language all three ethnic groups attained less well than majority pupils at the end of the 6th grade, but in terms of progress, ethnic groups demonstrated different patterns across subjects, as table 4.6 shows. In terms of final attainment score, Albanian pupils and pupils from other ethnic groups performed at a comparable low level. Although the attainment of pupils from the democracies of the former USSR was at a somewhat higher level, it was far behind the attainment of Greek pupils. A plausible reason why this occurred may be because many pupils who came from the democracies of the former USSR have a Greek origin, and therefore they may have members of their family who speak Greek. Pupils from the former USSR often arrive at the Greek school with some limited command of the Greek language.

Pupils from the former USSR were shown to have lower progress rates than majority pupils of Greek origin in language, ($p=0.003$), but not in *mathematics* where they were the only ethnic group that did not demonstrate less progress than the majority group ($p=0.874$). In mathematics, 'Albanians' and pupils from 'other' ethnic groups grouped in '*other*' category were shown to make less progress compared to the majority group at $p=0.026$ and $p=0.034$ respectively.

In language the lower progress rates for the former USSR ethnic group in relation to the majority group, associated with their lower final attainment scores in relation to their Greek counterparts may pinpoint the inadequacy of intervention provided to this ethnic group. Often schools do not provide support instruction to pupils from the former USSR, on the premise that to start with they were able to speak the majority (Greek) language to a certain extent, when they were enrolled for the first time in the Greek school. This finding indicates that pupils from the former USSR also need appropriate support teaching, especially in language, where they make less progress than the majority group. However for language, as has already been mentioned, some classrooms /schools were shown to be differentially effective in relation to the final attainment score of pupils from the former USSR at the end of year 6, as table 4. 10. (*‘Random effects in absolute attainment models – language’*), shows, findings that may indicate that some schools may have provided pupils from this ethnic group with support teaching, whereas other schools did not.

Table 4. 6. Examining the impact of belonging to a foreign/repatriated ethnic group on pupils' final attainment scores and progress rates for both subjects.

Final attainment score (in normal score units)	Coefficients & Standard errors (Standard errors are in brackets)			
	Attainment models		Progress models	
Models:	Mathematics	Language	Mathematics	Language
Model number:	17 th	17 th	12 th	14 th
p-value:	0.001	0.001	0.029	0.027
Constant	0.054 (0.042)	0.110 (0.036)	0.013 (0.028)	0.002 (0.020)
Initial score			0.810 (0.022)	0.915 (0.015)
Albanian	-0.473 (0.086)	-0.960 (0.084)	-0.118 (0.053) p = 0.026	-0.032 (0.037) p=0.396
The former USSR	-0.265 (0.141) p=0.06	-0.591 (0.136)	-0.013 (0.083) p = 0.874	-0.175 (0.060) p=0.003
Other	-0.637 (0.179)	-0.881 (0.169)	-0.256 (0.121) p = 0.034	0.031 (0.079) 0.69
Random terms:				
σ_{u0}^2 variance of the intercept at classroom level	0.093 (0.022)	0.055 (0.016)	0.044 (0.009)	0.023 (0.005)
σ_{u01} covariance term of initial score with the intercept			0.004 (0.005)	0.010 (0.003)
σ_{u1}^2 Classroom level variance of initial score			0.019 (0.006)	0.006 (0.003)
σ_{e0}^2 variance of the intercept at pupil level	0.863 (0.031)	0.849 (0.031)	0.279 (0.011)	0.141 (0.007)
σ_{e01} covariance term of initial score with intercept				0.024 (0.004)
σ_{e1}^2 variance of initial score at pupil level				0.025 (0.006)
-2loglikelihood	4389	4363.2	2519.9	1620.2
N (cases used)	1597	1608	1507	1503
All pupils were:	1858	1858	1858	1858
VPC at average initial score (%)	9.7	6	13.6	14

* Albanian pupils, pupils from the democracies of the former USSR and pupils from 'other' ethnic groups are compared to the final attainment scores of Greek pupils.

To compare the results between mathematics and language, Albanians and pupils allocated to the other ethnic group have lower attainment in relation to the majority group in both subjects. Pupils from the democracies of the former USSR attained marginally

less than the majority group in mathematics ($p=0.06$) but attained significantly less than the majority group in language (at $p<0.01$). In terms of progress, only pupils from the former USSR were found to progress less than the majority group in language, while Albanians and pupils from the the former USSR were found to progress less than the majority group in mathematics. However, the reader should be reminded that only 33 pupils belong to the 'Other' ethnic group and only 18 of them were tested twice, thus only these cases were employed for the estimation of the coefficients in the progress models. Although p-values for coefficients relating to a small ethnic group will tend to be large (reflecting the small sample size), thus making these coefficients not significant, it is worth being cautious; the large p-value may be due to having a small number of pupils in that group, and may become significant if a larger sample in the 'Other' ethnic group was employed.

An effectiveness criterion employed by Thomas and Collier (2004) is whether pupils who have limited proficiency in the school's language show gains in their achievement in relation to the majority group, thus progressing more in relation to the majority group. Greek primary schools do not meet this criterion of effectiveness since some ethnic groups, who speak Greek as a second language, make less progress in relation to the majority group in mathematics; while other ethnic groups make less progress in language.

According to Thomas and Collier (1997, p. 67) foreign/repatriated pupils *'must make more progress with each year of the school than the typical native speaker makes to ever close the academic achievement gap on school tests'*.

Given that they attain less than the majority group, in order that foreign/repatriated pupils may bridge their attainment gap with Greek pupils, they should demonstrate positive progress rates for subsequent years. Strand (1997) who examined attainment and progress differences between different pupil groups during Key Stage 1 in England, found that although pupils who spoke English as a second language had lower baseline attainment at entry, they showed positive progress rates in relation to their monolingual peers and consequently caught up with them at the end of Key Stage 1. Similar findings were reported by Schagen (1994) and Thomas (1995). When foreign/repatriated pupils are enrolled for the first time in primary Greek school on average their educational attainment lags behind the attainment of Greek pupils, mainly due to their limited competence in Greek (their second language). This limited competence negatively influences their comprehension in other subjects as well such as mathematics and science. Not only do foreign /repatriated pupils not demonstrate positive progress rates, but the opposite happens; certain ethnic groups make less progress than the majority group, thus falling further and further behind in relation to their Greek counterparts in certain subjects and other ethnic groups in other subjects. Then, provided that foreign / repatriated pupils make less progress in relation to their Greek counterparts for several school years, the attainment gap of foreign/repatriated pupils with Greek pupils is expected to widen instead of being reduced as pupils advance in secondary school. In subjects where foreign/repatriated pupils demonstrated no difference in progress in relation to their Greek counterparts, given that they had lower educational attainment than Greek pupils at the end of the 6th grade, it is expected that their attainment gap will remain unchanged, as Greek and foreign/repatriated pupil groups progress at the same rate from

one school year to the next. Such a subject is language, where only attainment differences were identified.

Different progress rates of ethnic minority groups have been established also in studies carried out in the UK. Strand (1998, p. 123, 129) in his analysis provides evidence of lower progress rates for certain groups of pupils, among whom were bilingual pupils needing English language support, and pupils entitled to free school meals, who usually come from lower social class backgrounds.

Absolute attainment models accounting for pupils' ethnicity point to the cumulative impact of ethnicity on pupils' final attainment scores at the end of primary school, while progress models adjusting for ethnicity show the *impact of ethnicity on pupils' progress rates* during the final year of primary school. The impact of ethnicity is greater in absolute attainment models since it has been shaped by their entire schooling experience. This experience is shaped by the impact of nursery school, the impact of mainstream classrooms, the general school impact (school climate), the impact of support schemes organised at the school level as well as by other socialising experiences offered by the wider community. However, in a progress model, adjusting for 'minority status' as well as for initial attainment score enables the estimation of the impact of having foreign/repatriated ethnic status on pupils' progress rates in relation to the progress made by Greek pupils. Yet, as variables adjusting for individual pupils' characteristics such as ethnicity, social class and gender are correlated with pupils' initial attainment score, controlling for pupils' initial attainment scores in a progress model partials out some of the variability associated with these individual characteristics. Hence the coefficients of minority status or belonging to particular ethnic groups, social class and

gender, when introduced in progress models, are smaller in magnitude than the coefficients of the same variables when introduced in absolute attainment models.

'It is acknowledged that because of the correlation between background factors and levels of initial attainment, the inclusion of initial attainment in the model may remove variance caused by the former' (Sammons et al., 1993, p. 38).

However, it is unrealistic to expect that foreign/repatriated pupils are able to achieve parity in their educational attainment with majority pupils as soon as they arrive in the host country's schools. Foreign/repatriated pupils should be expected ultimately to reach the same attainment level as their Greek counterparts, after the former have spent some years in primary school and after they have received some form of support instruction. The longer a foreign/repatriated pupil is exposed to the second language the better s/he appropriates the target language. The grade-level standards of the California Department of Education (1999) have spelled out accountability objectives for the schools to achieve in relation to minority pupils' performance. More specifically, they have said that *'immigrant students by their fifth year of enrolment in the United States will attain grade level performance for mathematics in the same proportion as mainstream pupils'* (California Department of Education 1999, p. 31). In other words a foreign pupil is expected to approach grade-related norms (the pupils' attainment is expected to lie near the 50th percentile of the attainment distribution derived through the administration of an authorised normative test to both native and foreign pupils) at the end of his/her 5th year of study in the host country's school.

Since in Greece there are discrepancies in the implementation of support teaching provision to foreign/repatriated pupils and given that the Ministerial decree Γ1/708/99 (published in ΦΕΚ 1789, Α part/ 28/9/1999) states that support teaching in reception classes should be provided to foreign/repatriated pupils for two years only, starting from the time they are for the first time enrolled in the Greek school, it is intriguing to discover whether foreign/repatriated pupils who have been enrolled for six or more years in a Greek school succeed in approaching grade norms comparable to those attained by Greek pupils.

In the above Ministerial decree it is assumed that after a two-year period, foreign/repatriated pupils should be able to approach the attainment level of their peers through exposure to both mainstream and support instruction.

For this purpose separate analyses were undertaken for foreign/repatriated pupils who, when tested, had already completed a) three years or more and b) six years or more including nursery education in the Greek school. The additional analysis shows that in fact the underachievement patterns remain.

a) After having been enrolled in the Greek school for three or more years, foreign and repatriated pupils still scored -0.408 normal score units less than the majority group in mathematics and -0.752 normal score units less than the majority group in language in terms of absolute attainment.

b) After having been in the Greek school for six or seven years, foreign/repatriated pupils still scored -0.368 normal score units less than the majority group in mathematics

and -0.667 normal score units less than the majority group in language in terms of absolute attainment.

For both mathematics and language it is noticeable that although the coefficient of being of foreign/repatriated ethnic status decreases in magnitude the more years a foreign/repatriated pupil has spent in Greek primary school, it decreases relatively little in absolute terms. The impact of having a foreign/repatriated ethnic background remains substantial and negatively influences even the attainment of foreign/repatriated pupils who have been enrolled in Greek schools from the beginning of their schooling. A possible reason for this striking pattern of underachievement is that foreign/repatriated pupils may have accumulated learning gaps in both subjects from their first years of enrolment in Greek primary school. These gaps remain even when they eventually understand the Greek language more. A possible reason that can lie behind these attainment and progress differences, as shown in section 4.5, which describes the implementation of support schemes, is that in many schools support teaching is not provided systematically to foreign/repatriated pupils.

4. 3. 2. *The Joint Effect of Ethnicity and Social Class*

Several studies have indicated that when factors associated with social deprivation are taken into consideration, attainment differences between majority and minority groups are diminished. In the UK, Tomlinson (1983) states that Mabey (1981) reported that poor environmental factors accounted for over half of the differences between black and white attainment in the ILEA literacy survey. The findings of current study also point to the interrelation of social class and ethnicity.

Tables 4.1 and 4.2 indicate that when an absolute attainment model adjusting for minority status also adjusts for pupils' social class, the coefficient of minority status is reduced by 38.3% and by 13.8% normal score units for mathematics and for language respectively. This reduction can be attributed to the over-representation of foreign/repatriated pupils in the lower social class category.

Ethnic minority pupils are disproportionately represented among pupils from lower social class background. This social phenomenon does not characterise only the Greek setting. Kumar (1994) in Goodwin (1995, p. 10) points out that in the U.K. the Swan committee (1985) was aware that the level of deprivation was considerably higher among minority ethnic groups.

In terms of progress for *mathematics*, minority status also reaches significance ($p=0.034$) when the progress model 3d (in table 4. 5.) additionally controls for social class. In this progress model both social class and ethnicity coefficients reach significance level. In the case of language a similar model could not fit.

In the UK Haque and Bell (2001, p. 366) created a multilevel model that examined differences in attainment between ethnic groups at the end of secondary school. These authors reported that when their models adjusted for variables measuring the social class construct such as the occupation of pupils' fathers and/or for the educational level of pupils' mothers, the existing differences in attainment or in progress between ethnic groups were substantially reduced. The authors suggested that variations in examination achievements among the minority ethnic groups could be partly understood via the effects of background variables, such as social class, instead of ethnic origin (p. 366).

Haque et al. (*op. cit.*) suggested that policy-makers, schools and teachers should differentiate these groups beyond their ethnic origin in order to reduce the nature of disadvantages for pupils from particular minority ethnic groups (p. 357).

Gilborn and Mizra (2000) have underlined the fact that pupils' ethnicity is very closely related to their social class position and that the effects associated with ethnicity and social class are interlinked. Ethnicity and social class have a joint effect on pupils' outcomes. Actually ethnicity and social class reflect different facets of educational disadvantage. The fact that ethnic differences are partly explained by social class variable and the fact that foreign and repatriated pupils are over-represented in low social class strata point to the necessity of initiating an educational policy that would be able to alleviate both conditions of disadvantage. In the Netherlands, the central goal of the Educational Priority Policy (EPP) was to reduce or eliminate the educational disadvantage of children insofar as this is a consequence of social, economic and cultural circumstances. The main target groups were non-indigenous and indigenous children whose parents have a low educational and occupational level (Driessen and Mulder, 1999).

Therefore, the impact of educational disadvantage should be alleviated through the adoption of a policy targeting both the negative impact of ethnicity and of social class.

The coefficients of having a minority status in absolute attainment and progress models adjusting for the effect of minority status only (as shown in the first square of the diagonal in tables 4.1 to 4.4) express the magnitude of these differences both in terms of attainment and in terms of progress. Subsequent arrangements to support foreign and repatriated pupils should be based on these facts.

However, these differences can be partially attributed to the impact of other variables, with which the ethnicity variable is correlated, such as social class. Coefficients of ethnicity in absolute attainment or progress models additionally controlling for social class point to the ‘net effect’ of ethnicity on pupils’ final attainment or on pupils’ progress rates, after controlling for the effect of pupils’ social class and after controlling for classroom membership (Sammons, 1995). In the current dataset this probably occurred because there was an over-representation of the foreign/repatriated pupils’ group in the manual and unskilled manual social class categories. The same pupils belong at the same time to the foreign/repatriated group and to the lower social class group. That fact causes the reduction of minority status coefficient when social class is introduced into the model.

Attainment differences associated with pupils’ ethnic status remain statistically significant (at $p < 0.001$) even when absolute attainment models additionally adjust for pupils’ gender and social class for both subjects, as shown in table 4.5. Progress differences associated with pupils’ ethnic status in mathematics remain statistically significant (at

$p=0.034$) when progress model 3d (in table 4.5.) additionally adjusts for pupils' social class. For language only significant progress differences (-0.165 , at $p<0.05$) for pupils from the democracies of the former Soviet Union were identified. This difference remains statistically significant (at $p<0.05$) even when the progress model additionally adjusts for pupils' gender, as model 4b (in table 4.5.) demonstrates.

To recapitulate, the claim that foreign/repatriated pupils underachieve in Greek primary schools can be based upon the following findings:

- a) The fact that for both subjects significant and substantial differences in attainment existed between minority and majority Greek pupils. These differences can also be identified between pupils from all three major ethnic groups and Greek pupils;
- b) The fact that these attainment differences can also be identified in analyses where only foreign/repatriated pupils with i) three or more years and ii) six or more years in primary Greek schools are concerned;
- c) The fact that significant attainment differences between foreign/repatriated and majority Greek pupils persist when the attainment models additionally control for pupils' social class and gender for both subjects;
- d) The fact that foreign/repatriated pupils and more specifically Albanian pupils and those from 'Other' ethnic background progressed significantly less than Greek pupils in mathematics;

- e) The fact that pupils from democracies of the former USSR progressed significantly less than Greek pupils for language.
- f) The fact that the previously mentioned significant progress differences between foreign/repatriated and Greek pupils continued to exist when the analyst additionally controlled for pupils' social class for mathematics.

All the above findings constitute evidence of underachievement for foreign / repatriated pupils.

4. 3. 3. The impact of social class on pupils' attainment and progress

Subsequently, the impact of belonging to different social class categories derived from the Goldthorpe social class stratification scheme will be considered. The social class variable was re-coded in order to be employed as an extended and dichotomous variable. Two baseline groups were defined: 'High grade professionals' serves as the baseline group in the detailed nine-category social class; whereas the five higher social class categories grouped together (higher classes) according to Goldthorpe social class classification scheme, served as base category in the new dichotomous social class variable.

Tables 4.1 and 4.3 show that when the dichotomous social class was introduced as a sole variable in absolute attainment models, pupils from lower social class background attained -0.412 and -0.510 normal score units for mathematics and for *language* respectively lower than pupils from higher social class. The coefficients of gender, social class and minority status retain statistical significance for both subjects when tested jointly in the absolute attainment models in table 4.5. Their coefficients in normal score units are

-0.207, -0.397 and -0.261 for mathematics and -0.528, -0.484 and -0.684 respectively for language. These coefficients are significant at $p < 0.001$.

The coefficient of social class also reduces slightly and becomes -0.316 for mathematics and -0.333 for language when the same absolute attainment model accounts also for the educational level of each pupil's mother in absolute attainment models. This is not a surprising outcome, since pupils' mothers belonging to a lower social class tend to have a lower educational level than the mothers of pupils belonging to a higher social class. Marmot et al. (1998) employed education as the main marker of socioeconomic position. In that sense a positive correlation should be expected between pupils' social class and the educational level of pupils' mothers. Indeed, the correlation between pupils' social class according to the detailed nine-category social class schema and their mothers' educational level is 0.512, significant at $p < 0.001$.

Gilborn and Gipps (1996), referring to a review of educational research on the achievements of ethnic minority children in the U.K. stress the importance of social class as a factor defining pupils' educational achievement over and above the influences of ethnicity and gender. According to these authors, '*social class is strongly associated with achievement regardless of gender and ethnic background: whatever the pupils' gender or ethnic origin, those from higher social class backgrounds do better on average*'. They also observe that '*When information on pupils' social class background is collected, there is usually a direct relationship with academic achievement; the higher the social class, the higher the achievement*' (Gilborn and Gipps, 1996, p. 16).

Gilborn and Gipps' comment holds for the pupils enrolled in primary schools in Piraeus. For mathematics, the representation of Greek low achieving³² pupils among social class categories gradually increases from 9.6% for pupils in the upper social class category to 27.2% for pupils in the lowest social class category. By contrast, for foreign/repatriated pupils the percentage of low achievers increases from 7.7% for pupils in the upper social class category to 47.1% for pupils in the lowest social class category. This descriptive data also indicates that there is a joint effect of social class with minority status. On the other hand, social class is strongly associated with achievement over and above the effects of gender and ethnic background. Whatever the pupils' gender or ethnic origin, those from higher social class backgrounds do better on average.

Next, the Goldthorpe social class schema (with nine categories) was introduced as a sole explanatory variable in an absolute attainment model predicting pupils' final attainment score a) for mathematics and b) for language. Each social class category was introduced as a dummy variable. The following social class coefficients were obtained according to absolute attainment models (11th):

³² Low achieving pupils are pupils who initially scored below the 27th percentile in all pupils' attainment distribution

Table 4.7. Examining the impact of social class on pupils' final attainment scores for mathematics and language (the detailed Goldthorpe social class scheme is employed).

Goldthorpe social class designation	Final attainment score (in normal score units)	Coefficients & Standard errors	
		Mathematics	Language
	Model number:	11th	11th
	p-value:	0.001	0.001
	Constant	0.374 0.154	0.494 0.151
I	High-grade professionals, administrators and officials; managers in large industrial establishments; large proprietors	Base category	
II	Lower-grade professionals, administrators and officials; higher grade technicians; managers in small industrial establishments; supervisors of non-manual employees	0.205 0.170	0.021 0.168
IIIa	Routine non manual employees, higher grade (administration and commerce)	-0.068 0.170	-0.135 0.167
IIIb	Routine non-manual employees, lower grade (sales and services)	-0.274 0.160	-0.417 0.157
Iva	Small proprietors, artisans, etc., with employees	-0.177 0.199	-0.232 0.200
Ivb	Small proprietors, artisans, etc., without employees	-0.445 0.169	-0.446 0.166
Ivc	Farmers and smallholders; other self-employed workers in primary production	The category was not found amongst this dataset	
V	Lower-grade technicians; supervisors of manual workers	-0.372 0.192	-0.379 0.191
VI	Skilled manual workers	-0.549 0.161	-0.722 0.158
VIIa	Semiskilled and unskilled manual workers (not in agriculture, etc.)	-0.665 0.168	-0.898 0.165
VIIb	Agricultural and other workers in primary production	The category was not found amongst this dataset	
	Random terms:		
	σ_{u0}^2 variance of the intercept at classroom level	0.072 0.019	0.048 0.016
	σ_{e0}^2 variance of the intercept at pupil level	0.824 0.033	0.825 0.033
	-2loglikelihood	3641.1	3633.6
	N (cases used)	1350	1353
	All pupils were:	1858	1858
	VPC	8	5.49

From the nine Goldthorpe social class categories, for which there was data in the current sample, the last four categories had a significant negative impact on pupils' final attainment scores for mathematics whereas five categories were found to have a significant negative impact on pupils' final attainment scores for language. For mathematics, the categories with significant coefficients were: *'proprietors not employing other personnel'*, *'supervisors / technicians'*, *'skilled manual workers'* and *'unskilled manual workers'*. For language, all the above categories were found to be significantly different from the base category plus one additional category, namely the *'Routine non manual employees, lower grade'* category.

In the case of language it is noticeable that non-adjacent social class categories turned out to be significantly different from the base category in terms of pupils' final attainment. Plausible explanations for this outcome can be attributed to the professional and financial structure of Greek society. Data from the EU-SILC study, which is conducted regularly by the National Statistical Service of Greece on behalf of the European Community aiming to assess conditions of poverty and disadvantage in each member state for the year 2003 reveals that skilled manual workers receive greater salaries than the two preceding occupational categories, namely *'small proprietors, artisans, etc., without employees'* and *'lower grade technicians; supervisors'*.

According to the 2nd dichotomous social class schema the last four social class categories were employed to form the category *'pupils from lower social class'*, so that the same detailed social class categories, derived from Goldthorpe scheme, are employed to form the category *'pupils from lower social class'* in both subjects. The initial five

categories, which in the mathematics analysis were not found to be significantly different from the Goldthorpe highest (base) category according to the detailed Goldthorpe scheme, were recoded to form the upper social class category.

The dichotomous social class variable was employed to form all interaction terms between social class, minority status, gender and initial attainment score, for simplicity reasons.

Pupils belonging to the lower (dichotomous) social class were expected to make less progress than pupils from higher social class in mathematics during year six, (coefficient -0.076 , significant at $p < 0.05$), but they are not expected to make less progress in language, as shown in the diagonal of tables 4.3 and 4.4. It can be inferred that since pupils from lower social class make less progress than pupils from upper social class in mathematics, if this trend persists over time, pupils from lower social class are prone to fall further and further behind in relation to pupils from upper social class as they progress in secondary school.

In England, Mortimore et al. (1988), Bondi (1991), Hutchison (1993) and Strand (1997) assessed the impact of social class on pupils' progress. All reported that pupils from disadvantaged backgrounds make less progress over the primary school years than their peers from more advantaged backgrounds.

In the USA Chall and Curtis (1991) state that the longer disadvantaged children stay at school, the greater becomes their attainment gap with the non-disadvantaged children.

Sammons (1995) reported that in the U.K. *'the gap in attainment associated with economic disadvantage and sex widened even further over the secondary school years'*.

'The ways in which social class affects educational opportunities are multiple and complex: some factors lie outside the school, others operate through institutional processes that disadvantage particular groups of pupils' (Gillborn et al., 2000, p. 19).

Therefore, this adverse social class effect leading to unequal opportunities for different pupil groups should raise awareness among educational agents in the direction of establishing, setting up and implementing support mechanisms dealing with educational and social disadvantage. These support mechanisms would increase educational and social opportunities for pupils from lower social class background who have lower educational attainment in both subjects and who have made less progress in mathematics in relation to pupils from higher social class.

4. 3. 4. Attainment and progress differences between boys and girls

For both subjects, gender has a negative impact on pupils' final attainment score, when introduced as a sole variable and also when introduced with additional explanatory variables. The magnitudes of gender coefficients are -0.245 for mathematics and -0.579 for language, both significant at $p < 0.01$, as shown in tables 4. 1. and 4. 2. The difference between girls and boys for language is more than double the difference between boys and girls for mathematics. In the case of language, this difference amounts to more than half a normal score unit. Boys perform less well than girls in their final attainment scores in both language and mathematics, but the difference is much greater for lan-

guage than it is for mathematics. In a report of 'recent research on gender and educational performance' in the U.K. (Arnot et al., 1998, p. 4) it was found that *'girls get off to a better start in reading than boys; the lead they have established by Key Stage 1 is maintained at Key Stages 2 and 3'*.

In terms of progress for mathematics, when gender is introduced as a sole variable in a progress model, it is of borderline significance, as table 4.3 shows. The coefficient of gender then is -0.054 (at $p=0.061$). That means that there is weak evidence that boys made slightly less progress than girls during the school year examined. For language, gender is significant (at $p<0.01$) when introduced without additional explanatory variables in a progress model. Its coefficient is -0.096 as table 4.4 shows. For language, there is strong evidence that boys make significantly less progress than girls. Given that significant differences in attainment between boys and girls in favour of girls exist for both subjects, if this pattern of boys making less progress than girls persists for several consecutive years, in other words if boys fail to maintain momentum with the girls, the attainment gap between boys and girls is expected to increase as pupils move in the lower grades of secondary school.

Similar results are derived from longitudinal British studies. In the School Matters study, girls progressed slightly more than boys between the ages 7 + and 10 + in inner London primary schools (Mortimore et al., 1988). In a review of '*recent research on gender and educational performance*' in the U.K. (Arnot et al., 1998, p. 19) it has been shown that *'girls have made somewhat greater academic progress between the ages of seven and sixteen than boys'*. In this study it is mentioned also that *'girls seem to do better in language across the world'*.

Kaluge (1998, p. 127) found that girls outperformed boys in terms of progress in a school effectiveness study carried out in Indonesia in both language and mathematics. In the Netherlands however, Bosker *et al.* (1990) reported that whereas differences in language attainment between boys and girls are very small, boys do better than girls in mathematics. In Cyprus, Kyriakides *et al.* (2000, p. 512) reported that in terms of progress boys achieve higher scores than girls in mathematics. This finding can be juxtaposed to the finding of the current study namely that Greek boys progress less than girls in mathematics; however this difference is of borderline significance.

Gillborn and Mirza (2000, p. 22) name the finding of boys' underachievement as the new gender gap as in earlier years girls used to have lower attainment than boys and not vice – versa. In a recent report of '*research on gender and educational performance*' in England (Arnot *et al.*, 1998, p. 7) it was found that '*boys and girls have recently been performing at very similar levels in Key Stage tests for mathematics*'. In the U.S.A. Gerber *et al.* (2001 p. 132) found that girls outperformed boys in both reading and mathematics from kindergarten through grade 3. But as Meighen (1997) states, '*a distortion is produced by looking at gender in isolation. Other variables, such as social class, ethnicity and region, are intertwined with gender*' According to Meighen gender differentiates the pupils after the age of 11. Gillborn and Mirza (2000, p. 22) report similar differences between the attainment of boys and girls at the end of secondary schooling (GCSE examination) in England. Arnot *et al.*, (1998 in Laura Sukhnandan 1999, p. 5) suggest that in England gender gaps are prominent in early literacy skills and later on in English, where generally girls outperform boys; this difference becomes

noticeable in GCSE exams, where girls make better progress overall between 11 and 16;

Gillborn and Gipps (1996) state that both social class and gender influence students' achievement and they should be jointly examined. Thus generalised statements about the superiority of girls over boys in attainment should be avoided, as in England, as well as in other countries differences between boys and girls are subject-specific.

'Blanket statements about girls performing better than boys or vice versa are difficult to justify; reference should always be made to a specific aspect of the curriculum' (Arnot et al., 1998, p. 8).

In the case of the current dataset, gender is significant, and its coefficient is not substantially reduced when the absolute attainment model additionally adjusts for pupils' social class and minority status along with gender, as models 6b3 for mathematics and for language show in table 4.5.

When the impact of gender is jointly examined with the impact of other individual pupils' characteristics for both subjects the coefficients of gender in absolute attainment models are reduced when the models additionally adjust for nursery school attendance or for hours of homework per week (as shown in tables 4.1 and 4.2). Thus, nursery school attendance and the weekly amount of homework undertaken by a pupil alleviate the impact of gender. The reduction of the coefficient of gender is more pronounced for mathematics. No other variable was found substantially to reduce the impact of gender.

Possible reasons lying behind boys' under-achievement according to the English National Literacy Trust (2003) are that boys may perceive studying as a girls' activity, and that examinations and tests may favour girls rather than boys. Also boys, being more disruptive, tend not to conform to what teachers say and also do not pay full attention to what is happening in the class, with the result that they learn less in the classroom and spend less time on the task. Also, boys belonging to a lower social class may have little incentive to perform academically, seeing that there are limited opportunities of success within the context of their families.

4. 3. 5. The Impact of Mothers' Educational Level

Mother's educational level has a significant positive impact on pupils' final attainment scores for both *mathematics and language*. The higher the educational level of a pupil's mother, the higher the pupil is expected to achieve at the end of year 6.

In most of the models mother's education has been used as an ordinal variable. Initially, mother's education was used as a dummy variable with six categories; it was found that its coefficients increased for every additional / higher educational level pupils' mothers have successfully completed. Subsequently the variable was used as an ordinal variable and it was found that pupils' expected final attainment score increased by 0.198 normal score units for mathematics and by 0.202 normal score units for language for every additional educational level pupils' mothers have successfully completed ($p < 0.01$). Next, a dichotomous variable for mother's education was created, where the impact of mothers having attained a higher than compulsory educational level was compared with the

impact of having attained education at a compulsory level only (nine years of schooling) or less.

In that case pupils with mothers who had completed a 'higher than compulsory' level of education were expected to perform 0.427 normal score units for mathematics and 0.450 normal score units for language higher than pupils whose mothers had completed a lower than compulsory educational level at the end of 6th grade. This difference in pupils' final attainment scores is quite substantial and amounts to almost half a normal score unit. Pupils with highly educated mothers achieve higher grades at the end of primary school. This outcome shows a quite deterministic trend that would persist unless the school acts in order to reduce the gap in educational attainment between educationally disadvantaged and educationally advantaged pupils.

The dichotomous variable 'mother's educational level' did not appear to have any significant impact on pupils' progress rates for both subjects. However, this variable had a significant impact on pupils' initial and final attainment scores. Maybe mother's education plays a role in giving the child a head start in his/her schooling, which is evident in pupils' attainment scores, but does not have an impact in terms of progress in later grades of primary school. The moderate impact of mother's education on pupils' final attainment scores is accounted for when the progress model controls for prior attainment.

4. 3. 6. The Impact of friends

The number of *friends* possessed by a pupil has a positive impact on his/her final attainment score and progress rates for both mathematics and language. The coefficient of friends when introduced on its own in an absolute attainment model is 0.148 for mathematics, while it is 0.139 for language, both significant at $p < 0.01$. The coefficient of friends when introduced on its own in a progress model is 0.033 for mathematics, while it is 0.013 for language, both significant at $p < 0.01$. These coefficients refer to the expected increase in a pupil's final attainment scores for one additional friend. Having friends seems to be a major integration factor, having a positive effect on all pupils' final attainment scores and progress rates while it also alleviates the impact that being a foreign pupil has on pupils' final attainment scores and on their progress rates. Pupils who have been named by more pupils in the class as 'a friend', or pupils enjoying a higher sociometric status, are expected to attain significantly higher scores and make more progress during the school year than pupils who have been named as 'a friend' by fewer pupils.

For language, the variable 'number of friends' remains significant when the progress model additionally adjusts for both gender and the former USSR ethnic group ($p = 0.0096$).

4. 3. 7. The Impact of Nursery School

Pupils who have been enrolled in *Nursery* school are expected to achieve higher final attainment scores in both subjects. Tables 4.1 and 4.2 show that these pupils are expected to achieve 0.379 normal score units for mathematics and 0.604 units for language higher than pupils, who have not been enrolled in nursery school (both coefficients are significant at $p < 0.01$). However, nursery school attendance does not seem to have any impact on pupils' progress during the 6th grade.

When the absolute attainment model jointly controls for nursery school and another explanatory variable, the coefficient of Nursery school takes its lowest value when the number of friends a pupil has has additionally been controlled for, as shown in tables 4. 1. and 4. 2. In this case there is common variation explained by Nursery school and friends, which might be attributed to the socialising experiences of the child in nursery school, which eventually end up with the child acquiring more friends.

When the absolute attainment model additionally controls for another explanatory variable along with Nursery school, then the coefficient of nursery school is also substantially reduced when the model additionally controls for pupils' ethnic status or for social class, as shown in tables 4. 1. and 4. 2.

The positive impact of nursery school in improving pupils' educational attainment has been acknowledged in the USA and in Britain. According to the American Department of Education (1998, p. 5) American districts can alleviate the impact of poverty through offering deprived children access to nursery school and early childhood programmes. According to this report, children from low-income families are about half as likely as

children from high-income families to attend pre-school programmes. Thus *'because there is a strong relationship between poverty, student achievement and low performing schools, districts can further focus their learning by intervening early to help children to be ready to learn'* (American Department of Education, *op. cit.*, p. 6).

In the U.K. Muijs and Reynolds (2001) in a review of the literature arrive at similar statements as well: *'That pre-school education does indeed have positive effects on students' subsequent achievement in primary school is demonstrated by research showing that children who take part in pre-school education perform better at primary school than students who have had no pre-school education'* (Muijs and Reynolds, 2001, p. 134).

In Britain the positive impact of pre-school provision on educational attainment in the case of minority pupils is highlighted by the Committee of inquiry into the Education of Children from Ethnic Minority Groups (Swann, 1985, p. 11), whereas in the case of deprived children it has been emphasised by the Plowden report (Central Advisory Council for Education, 1967). Hence the above reports suggested that the provision of nursery education should be given priority in disadvantaged areas and schools.

As an individual characteristic, nursery education has a positive effect at the pupil level in the U.K. according to a multilevel analysis carried out by Schagen (1994) at the end of Key Stage 1 in three out of four attainment targets (p. 167) in the core subjects of English, mathematics and science.

In addition, nursery school attendance has an alleviating effect on the impact of having a minority status, as already discussed in section 4. 3. 1.

4. 3. 8. The Impact of Homework

Pupils who devote more hours per week to doing their *homework* achieve higher final attainment score both for mathematics *and for language* than pupils who do less homework. Pupils are expected to perform 0.024 normal score units for mathematics and 0.022 normal score units for language higher at the end of the 6th grade for every additional hour they study during a ‘typical’ week. Homework makes a difference, as for 10 more hours of weekly homework pupils are expected to gain 0.240 normal score units for mathematics and 0.220 normal score units for language.

The impact of homework appeared to have a significant impact on pupils’ progress when introduced on its own in a progress model in mathematics ($p=0.04$). However, homework did not appear to influence pupils’ progress in language ($p=0.192$).

Recently, in the U.K. setting, two studies have examined the relationship between homework and achievement in primary schools, using a value-added approach (Tymms, 1997, Farrow et al., 1999). Both studies controlled for pupils’ initial attainment scores. Tymms (1997) found a positive relationship between hours spent on homework and pupils’ progress rates on a science test. In the Netherlands, Scheerens and Creemers (1992) showed that homework has a positive impact in raising educational attainment, while Van De Jong (1989) described homework as a causal mechanism boosting the effect of schooling.

When the absolute attainment model jointly controls for weekly hours of homework and a second explanatory variable, the coefficient of Homework decreases when the model additionally controls for pupils' social class (in the case of language only) or for the number of friends a pupil has (for both subjects). A plausible explanation for that effect is that pupils from higher social class have access to other social or cultural experiences that have the same effect (raising pupils' literacy level) as reading does. Also having friends gives pupils some opportunity to be exposed to language and use it as reading and writing does.

Holmes et al (1989, p. 36) underlined the link between homework and achievement and identified a tendency 'for boys whose parents are in middle class occupations to spend more time on homework than boys whose parents are in working class occupations' (Holmes et al., 1989, p. 40). The above authors disentangled a relation between homework and the social class of the child. They proclaimed that 'the interrelationship between parental schooling, time on homework and achievement falls into a virtually identical pattern to the interrelationship between parental class, time on homework and achievement' (Holmes et al., 1989, p. 42).

When in the same model as gender, homework slightly alleviates the impact of gender, as discussed in section 4. 3. 4.

4. 3. 9. Interaction terms in absolute attainment models

Aitkin and Zuzovsky (1994) stressed the danger of failing to include the interaction effects in multilevel models. Also as Grant and Sleeter (1986, cited in Banks, 1996) pointed out, much of the present research on multicultural education is limited in that it tends to examine the impact of race, class and / or gender instead of examining the interactions of these social constructs. If significant interaction terms can be included in a model, the impact of a variable taking part in an interaction term varies depending upon the value of the second variable forming this interaction term.

Table 4. 8. Examining interactions in absolute attainment models

Final attainment score (in normal score units)	Coefficients & Standard errors			
	Mathematics		Language	
Models:	With inter- action term 'foreign* absences'	With inter- action term 'foreign* friends'	With in- teraction term 'foreign* absences'	With in- teraction term 'foreign* friends'
Model number:	3d1a	3e1a	3d1a	3e1a
To be compared with:	3d1	3e1	3d1	3e1
p-value:	0.024*	0.0086**	0.034*	0.0015**
Constant	0.205 0.049	-0.560 0.059	0.267 0.045	-0.396 0.053
Minority status (1: foreign, 0: Greek)	-0.182 0.099	-0.443 0.102	-0.661 0.096	-0.923 0.102
Absences	-0.020 0.005		-0.014 0.005	
Interaction term 'Foreign * Absences'	-0.023 0.010		-0.021 0.010	
Friends		0.135 0.008		0.113 0.008
Interaction term 'foreign * friends'		0.066 0.025		0.080 0.025
Random terms:				
σ_{u0}^2 variance of the inter- cept at classroom level	0.081 0.022	0.126 0.025	0.055 0.017	0.070 0.017
σ_{e0}^2 Variance of the inter- cept at pupil level	0.857 0.036	0.701 0.025	0.828 0.034	0.715 0.026
-2loglikelihood	3412.7	4086.8	3353.8	4091.1
N (cases used)	1245	1597	1245	1600
All pupils were:	1858	1858	1858	1858
VPC (%)	8.6	15.2	6.2	8.9

According to absolute attainment models (3d1a) for both *mathematics and language* the *interaction term* between foreign/repatriated pupils and days they have been absent during the year is significant at $p < 0.05$. *Foreign pupils* suffered an additional negative effect on their final performance the more days they had been absent. For example, for mathematics, the coefficient of being a foreign/repatriated pupil (-0.182) increased by -0.023 for every additional day a foreign/repatriated pupil had been absent during the school year.

According to absolute attainment models (3e1a) for both *mathematics and language* another positive interaction term identified has been between '*minority status*' and *the number of friends a pupil has in the same classroom*'. The coefficients of these interaction terms are both significant at $p < 0.01$. The more friends a foreign/repatriated pupil has, the higher his/her predicted final attainment score becomes. Having more friends promotes the attainment of foreign/repatriated pupils more than it promotes the attainment of Greek pupils. For mathematics, the negative effect that being a foreign/repatriated pupil has on pupils' final educational attainment (-0.443) is alleviated by 0.066 for every additional friend a foreign/repatriated pupil has.

4. 3. 10. Interaction terms in progress models

Table 4. 9: Significant interaction effects in progress models

	Coefficients & Standard errors		
	Mathematics	Language	
Models:	With interaction term 'gender * initial score' ³³	With interaction term 'foreign * initial score'	With interaction term 'social class * initial score'
Model number:	3b3	3d2	9a
p- value	0.051	0.038*	0.005**
Constant	0.044 0.031	0.048 0.034	0.085 0.032
Initial score (in normal score units)	0.773 0.028	0.789 0.024	0.828 0.027
Minority status (1: foreigner, 0: Greek)	-0.105 0.043	-0.070 0.048	
Former USSR (1: from the former USSR, 0: Greek)			-0.179 0.063
Gender (1: boy, 0: girl)	-0.053 0.028		-0.085 0.023
Social class (1: lower, 0: higher)		-0.066 0.032	-0.042 0.027
'Male * initial score'	0.059 0.030		
'foreign * initial score'		0.098 0.047	
'Social class * initial score'			0.089 0.029
Random terms:			
σ_{u0}^2 Variance of the intercept at classroom level	0.044 0.010	0.049 0.011	0.023 0.005
σ_{u01} Covariance of the intercept with initial score	0.003 0.005	0.004 0.006	0.009 0.003
σ_{u1}^2 Variance of the slope (initial score)	0.020 0.006	0.019 0.006	0.005 0.003
σ_{e0}^2 Variance of the intercept at pupil level	0.277 0.011	0.279 0.011	0.134 0.007
σ_{e01} Covariance of intercept – initial score			0.021 0.004
σ_{e1}^2 variance of the slope (initial score)			0.024 0.006
-2loglikelihood	2515.2	2239.5	1381.4
N (cases used)	1507	1333	1326
All pupils were:	1858	1858	1828
VPC (%) at average initial score	13.7	14.9	14.6

For mathematics in terms of progress two interaction terms were identified:

The interaction term '*foreign by initial score*' has a significant coefficient (0.098, $p=0.038$) in progress model 3d2 only when the social class variable is included in the model together with the main effects minority status variable. This positive interaction term implies that foreign and repatriated pupils make more progress if they achieve highly at the beginning. In this case foreign/repatriated pupils' final attainment scores' gain for one additional normal score unit on their initial attainment scores is $0.789 + 0.098 = 0.887$ normal score units. In contrast, the gain Greek pupils make is equal to the coefficient of initial score, that is 0.789 normal score units only, according to this progress model (3d2).

The interaction term '*boys by initial score*' is marginally significant for mathematics ($p=0.051$), in model 3b3. Since the interaction term is positive, there is weak evidence that boys make more progress than girls for an additional normal score unit of initial attainment. The higher boys' initial attainment scores, the more they progress. According to progress model 3b3, although on average boys progress less than girls, boys' progress increases more for an extra unit of initial attainment score than girls' progress does.

For language, only one interaction term has been identified in terms of progress:

The interaction term '*social class by initial score*' is significant (at $p<0.01$) in model 9a. This can be explained in the following way: pupils belonging to lower social class made more progress if they achieved highly at the beginning of the school year.

33 This model cannot fit if it additionally controls for social class.

Other school effectiveness studies that identified significant interactions in terms of progress were those of Tizard et al. (1988) and Plewis (1991), who identified significant interactions between ethnic group and sex, and the study of Strand (1999) who identified significant interactions between ethnic group membership and free school meals and between ethnic group membership and sex.

4. 3. 11. Random coefficients in absolute attainment models

The existence of random coefficients in absolute attainment models is related either to different variance at level 1 associated with different pupil groups or to different variability in pupils' final attainment scores associated with different classroom settings (level 2) for different pupil groups. The latter should not be attributed to the conditions that occurred in these classroom settings within the school year investigated, but rather to the accumulated schooling experience of pupils enrolled in these classrooms. Whether different school and/or classroom environments contribute differently to the final attainment score of different pupil groups enrolled in them can only be addressed through the examination of random terms fitted in progress models.

Table 4. 10. Random effects in absolute attainment models

Final attainment score (in normal score units)	Mathematics	Language
Models:	With <i>minority status</i> random at pupil level	With ethnic groups Random terms with former USSR ethnic group at level 2 are identified
p-value	0.025*	0.011*
Constant	0.058 0.042	0.108 0.037
<i>Minority status</i> (1: foreign, 0: Greek)	-0.469 0.075	
Albanian		-0.950 0.084
Former USSR		-0.668 0.117
Other		-0.865 0.169
<u>Random terms:</u>		
σ_{u0}^2 variance of the intercept at classroom level	0.092 0.022	0.061 0.017
σ_{u01} Covariance of the former USSR ethnic group with intercept		-0.176 0.044
σ_{u1}^2 Variance of the former USSR ethnic group at the classroom level		0.270 0.145
σ_{e0}^2 Variance of the intercept at pupil level	0.832 0.033	0.849 0.031
σ_{e01} Covariance of ethnic group with intercept	0.108 0.054	
σ_{e1}^2 Variance of ethnic group at pupil level	0 0	
-2loglikelihood	4382.1	4354.1
N (cases used)	1597	1608
All pupils were:	1858	1858
VPC (%)	9.95	6.6

For mathematics, one significant random covariance associated with being a foreign/repatriated pupil at level 1 (pupil level) was detected, significant at $p=0.025$ in table 4.10. There is more variability among the final attainment scores of foreign/repatriated pupils in relation to Greek pupils for mathematics. Variance at the pupil level is 1.048 normal score units for foreign/repatriated pupils, whereas it is 0.832 normal score units for Greek pupils. This statement does not hold for language.

In addition, no significant random term at level 2 associated with pupils being of a foreign/ repatriated ethnic background could be fitted into absolute attainment models for mathematics. The absence of random terms associated with pupils' ethnic status at the classroom level implies that there is no evidence that classrooms are differentially effective for foreign/repatriated pupils.

For language, only a random coefficient at level 2 associated with belonging to the former USSR ethnic group was identified, significant at $p=0.011$ in absolute attainment models as shown in table 4.10. No other random coefficient could be fitted. In the long run, some classroom/school settings had a differential impact on the final attainment scores of pupils from the former USSR. Some classroom/school settings promoted the final attainment scores of pupils from the former Soviet Union more than other settings.

Possible reasons lying behind this finding may be that some schools assist pupils from the former Soviet Union by placing them in functioning support schemes, whereas other schools do not, on the premise that being repatriated, they already

have some knowledge of the Greek language. This differential treatment may lie behind this difference in their final attainment scores in language, created during the entire period of their schooling in Greek primary schools.

4. 3. 12. Random coefficients in progress models

The existence of significant random coefficients in progress models is associated with differential classroom effects, that is to say whether classrooms contribute differently to the progress of different pupil groups enrolled in them. In this study only random coefficients related to pupils' prior attainment were identified in '*progress model 2*' for mathematics as well as in '*progress model 4*' for language. No other random coefficients related to any other individual characteristics' variables were identified either for mathematics or for language.

4. 4. Contextual Effects

Contextual effects measure the effects of the school or the classroom intake composition on pupils' progress rates in each classroom or school. Since school effects were not identified in the current study, contextual effects measure the impact of compositional characteristics of each classroom on the progress rates of every pupil in this classroom.

These are over and above the main effects of individual pupils' characteristics. For example, pupils enrolled in classrooms where the majority of pupils are low achievers may influence their academic progress in a negative way. Similarly, pupils enrolled in classrooms where the majority of pupils enrolled are high achievers may benefit from affiliating in this environment in various ways (through collaborative learning, peer-group interaction etc.).

'Such research has been a feature of many British studies which have investigated the impact of concentrations of pupils from low social class, ethnic minority pupils and low average performance at entry (e.g. Nuttall, 1990; Sammons, et al., 1994; Sammons, 1996; cited in Opdenakker & Van Damme, 2001, p. 408).

The available research is conclusive with respect to the importance of the level of the intellectual and socio-economic composition of a school or a class: individual academic achievement is positively related to these characteristics of a class or a school.

Teddlie and Reynolds (2000, p. 184) state that *'the SES makeup of a school has a substantial effect upon students' outcomes beyond the effect associated with students' individual ability and social class'*.

The following variables measuring classroom contextual characteristics were introduced in 'progress models' to test the hypothesis that classrooms' contextual effects have an impact on pupils' progress. All the following variables were aggregated at the classroom level.

1. Pupils' average initial attainment scores
2. The variance between all pupils' initial scores in the class
3. The percentage of foreign/repatriated pupils
4. The number of foreign/repatriated pupils
5. The percentage of low achievers
6. The number of low achievers
7. The percentage of high achievers
8. The number of high achievers

9. The average amount of homework undertaken by the pupils
10. The average rate of absenteeism (average number of absences)
11. The gender composition in the class (percentage of boys)
12. The average educational level of the pupils' mothers
13. The percentage of pupils from lower social class
14. The class size

All the above characteristics were tested independently in two ways:

- a) In relation to a progress model adjusting for the main effects of the variables that were employed to estimate the contextual effects. For example, where the effect of the contextual variable 'percentage of foreign/repatriated pupils in the class' was tested, the main effects of 'minority status' had to be included in the progress model.
- b) In relation to progress models controlling for the related main effects corresponding to the contextual variables tested and additional individual pupils' characteristics.

For mathematics, contextual effects were tested compared to progress model 2, already controlling for pupils' initial attainment scores and also compared to the progress model additionally adjusting for 'minority status' and the interaction term 'minority status by initial attainment score'.

For language, contextual effects were also tested in relation to progress model 4. They were also tested in relation to a progress model adjusting for initial attainment score,

gender and pupils from the former Soviet Union ethnic group. They were also tested in relation to the progress model adjusting for initial attainment score, the dichotomous social class variable and the interaction term social class by initial attainment score.

Mortimore et al. (1988) make the distinction between school/classroom conditions classified as 'givens' and conditions shaped by the operation of the school or the classroom setting. From the above list the contextual characteristics 3, 4 and 11-14 can be characterised as 'givens' or as conditions not affected by the classroom operation, as these variables characterise a classroom setting before any pedagogical experience in the school takes place. For example, the classroom teacher cannot influence in any way the '*percentage of foreign pupils*', the '*gender composition*', the '*average educational level of pupils' mothers*', the '*percentage of pupils from low social class*' or the '*class size*'. The residuals derived from a progress model or from a contextual model adjusting for variables that can be considered as characteristics of a good teaching practice would be smaller compared to these classroom residuals derived from an unadjusted progress model. Hence, it is not right, based on residuals derived from a progress model controlling for classroom processes that define good pedagogical practice, to perform comparisons between classrooms for accountability purposes. The examination of classroom residuals derived from such models will result in *tautological reasoning*.

For example, amount of homework undertaken in a typical week is positively related to pupils' final attainment scores and to their progress during the school year investigated. Therefore it is not right to compare classroom residuals derived from a contextual model adjusting for the average amount of homework undertaken by the pupils in the

class on the premise that such a model controls for practices determining pupils' progress in the class.

On the other hand, '*class size*' directly influences the amount of teaching time as well as individual attention and interaction a classroom teacher can afford to assign to each pupil. Although '*there is still no clear consensus about the extent to which classes of different sizes promote the learning of students*' (Goldstein & Blatchford, 1998 p. 255) research has shown that the greater the class size the less progress pupils are expected to make. Therefore adjusting for class size may make classroom comparisons fairer when the goal is individual classroom accountability.

Van De Jong's (1989) study showed that the correlation between the percentage of migrants in a school and the schools' verbal test scores was -0.74 , therefore he considered it to be necessary to partial out the effects of pupils' ethnic background from the influences schools exert on pupils' test scores.

Alexander et al. (1979) investigated the impact of schools' social class composition on pupils' academic outcomes. In schools with high average SES levels, rather than low, any individual student is more likely to establish friendships with high SES classmates and this has a positive effect on his/her performance. Alexander et al. (*op. cit.*) reviewed studies undertaken by Duncan, Haller and Portes, 1968; Haller and Betterworth, 1960; Herriott, 1963; Kandel and Lesser, 1969; Kelly, 1952, Sewell, Haller and Ohlen-dorf, 1970; Simpson, 1962; Woelfel and Haller, 1971, Campbell and Alexander, 1965, Alexander and Eckland, 1975; Alwin and Otto, 1977; and Hauser, Sewell and Alwin, 1976, to establish that '*through well documented processes of interpersonal and refer-*

ence group influences the enhanced likelihood of entering into close relationships with high SES peers in schools of high average SES level is thought to yield numerous educational benefits' (p. 223). Thus this framework posits the interpersonal mediation of school average SES effects through networks of informal association between pupils.

Carpenter and Hayden (1987, p. 165) found that the sex composition of a school predicts a girl's exposure to key social influences, mediates the effects of the social structure upon achievement and plays a role in reinforcing the social and cultural advantages of particular girls. Bone (1983) and Steedman (1983) examined research findings relating to two specific claims, namely that girls in girls' schools get higher test scores and that girls in girls' schools choose subjects and careers that do not conform to feminine stereotypes.

Table 4. 11. Contextual models

	Mathematics	Language		
Models:	With Minority status, Minority status by initial score and % of pupils who are foreign and low - achievers	With former USSR, gender, social class, social class by initial score and % of low achievers	With former USSR, gender, social class, social class by initial score and % of high achievers	With former USSR, gender, social class, social class by initial score and cross level interaction % of high achievers by initial attainment score
Model number	3 cont 1	cont3	cont4	Cont cross
p-value:	0.035	0.01*	0.01*	0.038*
Constant	0.069 0.038	0.178 0.048	-0.066 0.043	-0.094 0.046
Initial score (in normal score units)	0.796 0.023	0.828 0.027	0.820 0.027	0.753 0.041
Minority Status (1: foreign, 0: Greek)	-0.086 0.045			
Former USSR (1: from the former USSR, 0: Greek)		-0.175 0.063	-0.176 0.063	-0.177 0.063
Gender (1: boy, 0: girl)		-0.085 0.022	-0.089 0.023	-0.087 0.023
Interaction term 'Minority status by initial attainment score'	0.071 0.044			
Dichotomous social class 1: lower social class 0: higher social class		-0.036 0.027	-0.040 0.027	-0.038 0.027
Interaction term 'social class by initial attainment'		0.081 0.030	0.088 0.029	0.094 0.029
Contextual effects				
% of foreign and low achievers	-0.986 0.488			
% Low achievers		-0.385 0.149		
% High achievers			0.595 0.125	0.681 0.132
% High achievers by initial score				0.236 0.110
Random terms:				
σ_{u0}^2 Variance of the intercept at classroom level	0.041 0.009	0.021 0.005	0.014 0.004	0.014 0.004
σ_{u01} Covariance of the intercept with initial score	0.005 0.005	0.010 0.003	0.007 0.002	0.006 0.002
σ_{u1}^2 Variance of the slope (initial score)	0.020 0.006	0.005 0.003	0.005 0.003	0.004 0.003
σ_{e0}^2 Variance of the intercept at pupil level	0.278 0.011	0.133 0.007	0.135 0.007	0.135 0.007
σ_{e01} Covariance of the intercept with initial score		0.021 0.004	0.021 0.004	0.021 0.004
σ_{e1}^2 Variance of initial score at pupil level		0.025 0.007	0.022 0.006	0.023 0.006

-2loglikelihood	2515	1375.2	1363.1	1358.7
N (cases used)	1507	1326	1326	1326
All pupils were:	1858	1858	1858	1858
VPC (%) at average initial score	12.8	13.6	9.3	9.3

For mathematics, contextual model 3_cont_1 showed that *'the percentage of foreign/repatriated pupils who are low achievers in a mainstream class'* has a negative effect on the expected progress of all pupils enrolled in this class over and above the effects of pupils' prior attainment, their ethnic status and over the effects of the interaction term of 'ethnic status by initial attainment score'. In classrooms with a large percentage of foreign/repatriated pupils who are low achievers, a negative contextual effect is reducing all pupils' performance. Based on this finding one can argue that providing support to initially low achieving foreign and repatriated pupils is expected to raise the educational outcomes of majority pupils as well. It should be noted that a similarly specified contextual model additionally controlling for social class (along with minority status, and the interaction term of ethnicity by initial score) could not fit ($p=0.083$).

No other contextual effects could be identified at $p<0.05$ for mathematics.

For language, the variable 'percentage of low achievers in the class' was significant when introduced in the contextual model cont3 at $p=0.01$. In addition, the contextual variable 'percentage of high achievers in the class' was significant at $p=0.01$, when introduced in a contextual model cont4. Both the above models cont3 and cont4 already control for initial attainment score, pupils' gender, former USSR status, social class and the interaction term social class by initial attainment score. The two contextual terms can be interpreted in the following way: The percentage of low achievers in a class has

a negative impact on the expected progress of all pupils enrolled in this class. The percentage of high achievers in a class has a positive impact on the expected progress of all pupils enrolled in this class.

Levine and Lezotte (1995, p. 532) suggested that ‘the number of low achievers in a class affects teachers’ decisions regarding pacing of instruction, their ability to respond appropriately and provide effective assistance to students, and overall behavioural dynamics in a classroom (Barr & Dreeben, 1983; Dreeben & Barr, 1988a, 1988b; Leinhardt & Pally, 1982; Levine, 1985a)’.

For language, only one cross-level interaction term could be fitted. The model `cont_cross`, which includes the cross level interaction ‘percentage of high achievers in the class by initial attainment score’, was a significant improvement on the contextual model `cont4` ($p=0.038$).

The later cross level interaction effect implies that the positive impact of the contextual variable ‘*percentage of high achieving pupils in the class*’ increases as pupils’ initial attainment score increases. The higher the initial attainment of a pupil, the more his/her progress is enhanced by the existence of a ‘high percentage of high achieving pupils in the class’.

Given that the slopes of the classroom lines are modified if cross-level interaction effects exist, the impact that ‘pupils’ initial attainment scores’ on ‘pupils’ final attainment scores’ is not fixed, but depends on the ‘*percentage of high achieving pupils*’ in the class. The existence of the above cross-level interaction term implies that the presence of a ‘high percentage of initially high achieving pupils in a given class’ not only boosts

the expected final attainment scores of every pupil (change of intercept) in this class but also changes the slope associated with initial attainment score. This entails that initially high achieving pupils gain more from being in a classroom with a high percentage of high achievers, than low achievers do.

To summarise, for language the greater the percentage of pupils who initially have been low achievers in a class the more the performance of all pupils enrolled in this class is expected to decline. The higher the percentage of pupils who initially are high achievers in a class the more the performance of all pupils enrolled in that class is expected to be boosted, as positive contextual effects emerge. In addition, in classes with 'a high percentage of high achieving pupils', pupils who initially were high achievers progressed more than pupils who initially were low achievers. Classrooms, where a majority of highly achieving pupils exist, boost the performance of high achievers, more than they boost the performance of low achievers.

This finding resonates with Wang et al. (1995, cited in Borman and Rachuba, 2001, p. 3) who suggested that '*students who attend schools with high concentrations of under-achieving, poor, and minority peers may be placed at increased risk of academic failure*'. In classrooms with a high concentration of low achieving pupils a substantial negative contextual effect is expected to decrease the performance of all pupils.

4. 5. The implementation of support schemes

Given the negative impact of ethnicity on pupils' achievement and sometimes on pupils' progress, this section will deal with the degree to which support teaching to foreign/repatriated pupils has been implemented so far. This information came from principals' and support teachers' responses to questionnaires administered to all 58 schools visited. 44 principals replied to the question 'for how many months in the last school year were support schemes running in their school'.

The lower educational attainment for foreign/repatriated pupils and for certain ethnic groups the lower progress rates for some ethnic groups in mathematics and for some in language, which were derived from the multilevel analysis, can be related to the implementation of support schemes during the school year the study was conducted, which was not too different from previous years. The previous analysis shed light on the existence of different patterns of 'underachievement' between disadvantaged groups. Principals' and support class teachers' accounts confirm that support schemes have been insufficiently and inconsistently implemented throughout the schools investigated.

However, on the basis of the findings of minority pupils' underachievement no causal statements should be made about the inadequacy of support schemes as there is insufficient implementation of these schemes; *'A policy can only lead to the desired results, when the decisions taken by those pursuing the policy are actually put into practice'* (Driessen and Mulder, 1999, p. 53). Furthermore the insufficient implementation of these schemes is not the only reason why patterns of underachievement between foreign/repatriated pupils' outcomes in relation to majority pupils' outcomes exist. The

informed reader can think of many other plausible reasons, such as the pedagogical programme(s) adopted there, the climate and the instruction in the mainstream class, whether multicultural elements have been incorporated into the school's programme, low teachers' expectations, racism, etc. However, the findings of insufficient implementation provide one plausible explanation, which along with other explanations can account for the findings of foreign/repatriated pupils' underachievement. Furthermore, it aims to emphasise the schools' responsibility in assisting foreign/repatriated pupils with limited skills in Greek or with accumulated learning gaps, as well as Greek low achieving pupils, in catching up with the rest of the class, assisting them to integrate successfully in the mainstream.

The finding that foreign/repatriated pupils underachieve compared to Greek pupils show that the current support arrangements along with the whole schooling experience in the mainstream are not enough to close the initial attainment gap between foreign/repatriated pupils and majority pupils. The findings that pupils from lower social class underachieve compared to pupils from higher social class constitute *cause for concern*. These findings serve as a starting point for looking more deeply at issues associated with the provision of support teaching for foreign/repatriated pupils and for pupils from lower social class. Principals and support teachers suggested that support schemes for foreign/repatriated pupils are inadequately implemented in Greek state primary schools. No support schemes categorically targeting low achieving Greek pupils who in the majority come from a disadvantaged background exist at the moment in primary Greek schools. The absence of any support scheme targeting pupils from lower social class background can be related to the finding of lower attainment among them in

relation to pupils from higher social class background in both subjects. In addition, it is reflected in the finding that pupils from lower social class (disadvantaged backgrounds) make less progress in mathematics and for which reason their attainment is expected to fall further and further behind in relation to pupils from higher social class.

According to the principals' accounts, reception classes are less adequately implemented than coach classes as reception classes operate in parallel to the mainstream. Districts have to reappoint teachers to work as support teachers in schools where reception classes are operating, and this is a very time-consuming enterprise. Districts do not appoint teachers to work in reception classes until several months of the school year have already passed. Sometimes they do not appoint any teacher at all, in spite of the school's application. However, coach classes function more regularly, since teachers already appointed to a mainstream class teach in them in the afternoon on an overtime basis. Therefore, from an administrative point of view it is easier for coach classes to operate.

According to principals' accounts, from a total of 9 months of a typical school year, coach classes operated for an average of 7 months, whereas reception classes operated only for an average of 4.7 months.

However, according to support teachers' accounts in 6 schools, even when teachers were appointed to work in the reception classes, they often had to replace mainstream class teachers who were absent or on leave. As reception classes functioned for a limited time only, foreign/repatriated pupils have been exposed to piecemeal and fragmented support instruction.

Large variations in support teaching policies among schools have often been the rule rather than the exception, with negative consequences for equal educational opportunities for foreign/repatriated pupils and for Greek pupils from lower social class who underachieve.

In some cases, on the other hand, failure to create a coach class in a school was attributed to the lack of interest on the part of mainstream class teachers in working extra hours.

Under these operational constraints, the phenomenon of lower educational attainment for foreign/repatriated pupils in comparison to the majority group should not be attributed to the ineffectiveness of current support schemes. It could be attributed at least partly to the fact that not all schools with minority pupils apply for support schemes categorically targeting these pupils and to the limited implementation of these support schemes (the fact that support schemes operate for less than nine months in the majority of schools involved).

Kavouri (1996, p. 142) claims that the success of an innovation does not depend only on the characteristics of the innovation itself. Reception classes and Coach classes can be perceived as an educational innovation, since they have been initiated relatively recently through law N1404/83, n^o 45. The implementation of Reception classes and Coach classes had been scanty up to the year the current study was conducted, as it was up to principals' discretion to apply for a support scheme, while it was up to district administrators' discretion to appoint extra teachers.

4. 6. The importance of the provision of effective learning time

Augmented effective teaching time or the effective learning time is a component identified in most successful³⁴ support teaching programmes, such as Success for All (Slavin et al., 1996) and Reading Recovery (Pinnel, 2000), already described in sections 1.4.2. and 1.4.3. More effective teaching time may stem from either more direct instruction in the mainstream, or through the provision of extra teaching time after the end of the school day. The total teaching time remains the same if foreign/ repatriated pupils are supported in a pullout setting (in Greece in a reception class that operates in parallel to the mainstream class). Attendance in a reception class may increase the effective instructional time for foreign/repatriated pupils, as in the reception class the level of instruction is more adapted to actual level of these pupils, than in the mainstream class.

Foreign/repatriated pupils with limited competence in Greek may not be able to comprehend large parts of instruction in the mainstream class. If support is provided, foreign/ repatriated pupils can be assisted either to raise their level in the Greek language so that they can understand both the support and the mainstream instruction better, or to better comprehend the curriculum content taught in the mainstream class. Support schemes, by clarifying and consolidating aspects of the curriculum

34 Programmes that provide intervention for disadvantaged groups such as Success for All and Reading Recovery have been reported to have substantial effect sizes. The effect sizes reported for Reading Recovery were +0.35 and +0.75 for dictation and text reading level for Reading Recovery (Pinnel, 2000). The effect sizes reported for the lower quartile of students enrolled in Success for All were particularly positive, ranging from ES = +1.03 in the first grade to ES = + 1.68 in the fourth grade' (Slavin and Leighton, 1990).

that have been taught in the mainstream class, make a greater part of the mainstream class lesson comprehensible.

Ultimately, the time foreign and repatriated pupils spend in the mainstream becomes effective learning time, as they can understand the content taught in the mainstream class better.

If foreign/repatriated pupils can understand the lessons taught in the mainstream class, this can enhance their linguistic capacity in the second (Greek) language through acquisition. Krashen (1990) confirmed that second language learners need access to comprehensible input to be able to acquire a second language. Minority pupils need instruction tailored to their level in order to become capable of profiting to the fullest extent from what has been taught to them. Minority pupils not enrolled in a support instruction programme comprehend only part of the curriculum taught in the mainstream class, and hence they take longer to acquire the second language and approach grade norms.

Krashen claimed that *'an individual acquires language through exposure to language that is understandable, yet contains new grammatical structures just beyond the learner's current level'* (Krashen, 1981 in Smith & Heflin, 1988, p. 11).

Krashen and Terrel (1988, p. 1) stated that pupils can acquire a second language by understanding messages, that are a little beyond the current level of pupils' language competence. This is the $i+1$ or comprehensible input hypothesis. Sometimes the level of the lesson in the mainstream class is too advanced for them to understand, and therefore these pupils need to be assisted in a support class. Thus foreign pupils can obtain com-

prehensible input in both the mainstream and in the support class. This can be achieved if the lesson delivered in the mainstream and the support class is coupled with several materials and visual aids and if the teacher's speech is modified to include confirmation and comprehension checks, clarification requests, and repetitions. The support class lesson can back up the mainstream class lesson by repeating and explaining key concepts taught during language and the mathematics mainstream lessons, or teach prerequisite concepts to these pupils in a language level closer to their actual language level. In the support class it is easier to supply foreign/repatriated pupils with comprehensible input, since the support class teacher can focus only on their needs and can plan the support lesson accordingly. Krashen (1982, p. 66) states that optimal input should be comprehensible, interesting and relevant, and not grammatically sequenced. Finally, optimal input should be in sufficient quantity for acquisition to take place. Hence, a support class can provide this additional input and the additional exposure time required.

It is important that the minority pupils with limited competence in Greek be able to gain from both being exposed to mainstream class instruction and attending a support class, as minority pupils enrolled only in the mainstream class do not have access to the curriculum taught. Foreign/repatriated pupils may be physically present in the mainstream classroom, but they are engaged in the teaching that takes place for a limited time only, as they partially understand what is being taught. Being able to understand the content of instruction makes all the difference in respect of them acquiring the second (Greek) language and increasing their academic performance.

The crucial issue for these pupils' learning in the mainstream, as well as in the support class, is *'how can the proportion of time that foreign and repatriated pupils are truly engaged in learning be maximised?'*

Minority pupils need less time to approach grade norms when provided with support instruction in relation to the time needed when not provided with support instruction. Given that effective support teaching programmes such as Slavin et al.'s (1990) 'Success for All' and Pinnel's (1988, 2000) 'Reading Recovery' were reported to have substantial effect sizes, they bring the supported pupils closer and earlier to the average in relation to unsupported pupils.

However, the time needed to approach grade norms is relatively long: According to Cummins (1984) *'5-7 years are required on average to immigrant students who came to the country after the age of six, to approach grade norms in academically related aspects of the second language proficiency'*. Hakuta (2000) proved that it takes the pupil between four and seven years to attain academic proficiency in the second language. This period may be shorter or longer according to each child's particular circumstances. Therefore foreign and repatriated pupils should follow a support-teaching programme for as long as their performance in basic skills lags behind the average performance of Greek pupils.

5. CONCLUSIONS

5. 1. The absence of school effect

This study clearly indicates the absence of school effects among the Greek state primary schools in the wider area of Piraeus during the final year of primary school in language and in mathematics. The absence of school effects was found in a variety of null, absolute attainment, progress, adjusted-progress and contextual models. However, more school effectiveness studies are needed that would be based on a sample comprising broader socio-economic areas, or comprising urban and rural areas in order to identify whether school effects exist. Furthermore, longitudinal studies based on an extended time span including several school years may reveal school effects that could not emerge, given the restricted time span of six months that elapsed between the initial and final assessment carried out in this study.

On the other hand, the current study revealed the existence of significant classroom effects. These effects were present according to a variety of null, absolute attainment, progress, adjusted progress and contextual models. Hence it can be concluded that classrooms matter and make a difference to pupils' progress. The above findings are in line with the findings from many previous international SER studies, many of which were carried out in the Netherlands and in Australia indicating that classroom effects are greater than school effects in explaining students' progress (Creemers, 1994; Reynolds et al., 1996, Scheerens and Bosker 1997, Hill and Rowe, 1996).

5. 2. Conditions associated with educational disadvantage

The current study has shown that foreign and repatriated pupils underachieve in relation to their Greek counterparts. The findings are more disquieting for mathematics, where both Albanian pupils and pupils from 'other' ethnic group make less progress in relation to the majority group, whereas only the group of pupils originating from the democracies of the former Soviet Union was shown to make less progress in relation to the majority group in language (as shown in table 4.6.). In addition, pupils from lower social class were found to underachieve in relation to pupils from higher social class; pupils from lower social class were also shown to make less progress in relation to pupils from higher social class in mathematics (table 4.5.), whereas for language social class was found to interact with initial attainment score in terms of progress (table 4.9.). In addition, boys were found to underachieve in relation to girls in both subjects, but boys were found to make less progress in relation to girls only in language (tables 4.5. and 4.9.). Furthermore, substantial attainment differences between foreign/repatriated and majority pupils, between pupils from lower social class and pupils from higher social class and between boys and girls existed for both subjects. The above findings are in line with the English, American and Dutch literature on underachievement of ethnic minority pupils and of pupils from disadvantaged backgrounds (Swann, 1985; Gillborn and Gipps, 1996; Crawford, 1997; Eccels, 1997; United States Department of Education, 1994; Olne, 1995, Van De Jong, 1989). Studies with similar findings carried out within the school effectiveness paradigm in the UK were those by Mortimore et al. (1988), Bondi (1991), Hutchison (1993) and Strand (1997, 1998, 1999)).

If underachieving groups do not demonstrate positive progress rates in relation to their comparison groups in the future, they are not expected to catch up; The performance of ethnic groups that were shown to make less progress in relation to the majority group is likely to fall further behind the performance level of Greek pupils during lower secondary education (Gymnasion) in subjects where progress differences have been identified. Since pupils from lower social class were shown to make less progress in relation to pupils from higher social class in mathematics, their performance is expected to fall further and further behind the performance of pupils from higher social class during secondary school. Attainment differences are expected to remain intact for groups and subjects where no progress differences have been identified. The lack of support schemes for low achievers and foreign/repatriated pupils during lower and upper secondary education in Greece (in Gymnasia and Lykeia) and the increasing complexity of the curriculum in secondary education may make it more difficult to counteract attainment differences between pupil groups who have initially fallen behind and the majority group.

5. 2. a. Implications stemming from the underachievement of foreign/repatriated pupils

These substantial attainment and progress discrepancies between foreign and repatriated pupils and Greek pupils in language and in mathematics indicate that they need to be supported in a more coherent, intensive and systematic way in both subjects. The biggest difference in final attainment scores is between foreign/repatriated pupils and majority pupils in language. This difference amounts to almost one normal score unit.

Lower progress rates for pupils from the former USSR in language and of Albanian and of pupils belonging to the ‘other’ ethnic group in mathematics suggest that all these three ethnic groups need prompt and intensive support in order to raise their attainment in both subjects by creating positive progress rates. The situation needs more urgent monitoring for ethnic groups and subjects where lower progress rates have been identified. It is crucial that foreign/repatriated pupils should be supported not only in language but in mathematics as well, in order to cover their learning gaps and help them to understand the content of the mathematics lesson. Under the current support teaching arrangements, foreign/repatriated pupils are assisted in mathematics in the ‘*support with homework*’ sections only (in schools where such sections operate), along with the Greek counterparts; they are not supported in a separate setting. Targeting foreign/repatriated pupils with limited competence in Greek separately as a group is required in order to help them acquire the vocabulary that is indispensable to understanding concepts in mathematics and at the same time understand the requirements and the contents of the mathematics lesson. If these lower progress rates characterise these ethnic groups during secondary education, their educational attainment is bound to fall further and further behind the attainment of their Greek counterparts.

5. 2. b. Implications stemming from the underachievement of pupils from lower social class

In addition, the finding of pupils' lower educational attainment in both subjects and also the finding of lower progress rates of pupils from lower social class for mathematics show that support teaching arrangements have to be additionally made for low achieving pupils in general, as these pupils frequently originate from a lower social class background. Higher priority should be given to the support of low achieving pupils from lower social class in mathematics, as these pupils were found to have lower progress rates.

5. 3. The gender issue

Boys attained less well than girls in both mathematics and language lessons and made less progress than girls did in language. If differences in progress between boys and girls are not properly addressed, either with satisfactory support teaching arrangements or in the mainstream class setting, not only are these attainment differences bound to continue, but they are expected to be exacerbated during secondary school years, as boys are likely to fall further and further behind girls during secondary school. The reasons why boys underachieve are different from the reasons related to the underachievement of ethnic minority pupils and of pupils from low social class and therefore separate studies are needed.

5. 4. Ways to counteract educational disadvantage

The complex picture of underachievement described above entails that affirmative steps should be taken in order to alleviate these educational and social inequalities. These affirmative steps can take the form of support instruction in a separate setting, the support classroom or in the mainstream classroom. A comprehensive framework of intervention, catering for the educational needs of every pupil should be created, as such a comprehensive support teaching venture has the potential to deal effectively with attainment and progress discrepancies associated with educational and social disadvantage, or in other words in an attempt to partially alleviate the impact of ethnicity and of social class.

This framework has to provide intensive instruction in the Greek language in every school in which foreign and repatriated pupils with limited competence in the Greek language are enrolled; in addition, it has to provide additional support schemes to assist these pupils to comprehend the curriculum covered in the mainstream class. This framework also has to provide for the needs of low achieving Greek and foreign pupils who have attained a satisfactory level in Greek language. Low achieving Greek pupils who cannot be characterised as pupils with special educational needs but who in the majority come from a lower social class background need also to be targeted in a support instruction programme along with minority pupils who have already acquired the Greek language at a satisfactory level. These two pupil groups can be taught together other lessons such as mathematics or they can be taught language at a higher level, if their learning needs and/or learning gaps coincide. Slavin et al. (1990) stressed the need

for regular assessment exercises in order to create homogeneous groups of pupils, with similar learning needs.

Support instruction should be more widely implemented among schools and districts, within a framework of accountability and support for classrooms, schools and school districts. Support instruction has to address the needs of every foreign/repatriated pupil in need of support, and of every pupil who has fallen behind in the two basic subjects even in schools where teachers appointed are not interested in teaching on an overtime basis in an after-school support scheme. This comprehensive framework can either be conceptualised within the current educational framework, or with minor additions to the current framework.

This framework should legislate for the provision of support instruction categorically targeting two pupil groups: a) foreign / repatriated pupils with limited competence in Greek and b) low achieving Greek pupils, in order to counteract and alleviate educational and social inequalities.

Coach classes and reception classes were initially legislated for in N1404/83 no 45 in order to cater for the needs of foreign and repatriated pupils. Among other aims, coach classes and reception classes aimed to teach Greek as a second language to foreign/repatriated pupils and/or to assist them to carry out their homework assigned in the mainstream class. However, they are often abolished in practice; in many schools they are not initiated; in most others they are very poorly implemented; in others they are replaced by the 'whole day school', recently established by law 2525/97 (ΦΕΚ 188 τ.Α) and the accompanying circular Φ13.1/897/ Γ1/694/6-9-2000. Although the legislating

framework regulating the creation of ‘whole-day schools’ was a very positive step as it catered for the needs of pupils with working parents providing them with support for homework and enrichment programmes, it does not categorically cater for the needs of foreign/repatriated pupils with limited competence in Greek or for the needs of low achieving Greek pupils. The whole day school with certain legislative amendments might prove to be a positive initiative for disadvantaged pupils, since the more time pupils spend in school, the more their attainment is raised and the more the impact of variables associated with disadvantage is alleviated. Findings from the international School Effectiveness Research Project (Reynolds et al., 2002, p. 277) suggest that *‘pupils’ increasing time in school weakens the relationship between pupils’ achievement and their parental, ethnic, educational and social class backgrounds’*.

However, legislators and districts have to think hard about equity, as under the current framework equal time is allocated to high and low achieving pupils and to minority and majority pupils in order to support them with homework in any given school that has adopted the whole day school programme.

It is not rare for only one or two ‘support with homework’ sections to cater for the needs of all pupils enrolled in classes 1-6 of primary schools. In these cases, since no categorical provision has been made for the needs of low achieving pupils or of foreign/repatriated pupils with limited competence in Greek, it is unlikely that most of the support time available would be allocated to low achieving pupils; besides, hardly any time is allocated to teaching Greek to foreign and repatriated pupils, whereas the time allocated to making the lesson of the mainstream class comprehensible to them is insufficient. Furthermore in the framework of the whole-day school there is no provision for

diagnostic assessment or grouping of low achieving pupils according to their learning needs.

Law 2525/97 explicitly states in article 4, (3rd paragraph) that pupils with learning difficulties have priority in being registered in ‘whole day school’ sections or in ‘sections of extended timetable’. Yet, in practice, support with homework is equally provided to every pupil who is registered in whole day school, on the premise that such a provision would raise all pupils’ attainment. Since in the same setting high and low achievers are supported with homework and allocated exactly the same time, this arrangement is insufficient to address educational inequalities. The attainment gap between minority and majority as well as low achieving and high achieving pupils is expected to remain intact, as disadvantaged pupils are neither provided with the extra time required to catch up with their advantaged peers, nor with instruction tailored to their level.

An example of how such a categorical support can be provided to disadvantaged groups can be taken from the American ‘No child left behind Act’ (United States of America - Congress House, Boehner, 2001), which sets out a framework for allocating funds and services to disadvantaged pupil groups and to schools serving a majority of disadvantaged pupils. Within this framework funds for support services (Title I funds) can be used for general school purposes only when 40% or more of a school’s intake suffers from educational disadvantage in ‘*schoolwide programmes*’. In cases where less than 40 percent of the school’s intake suffers from educational disadvantage, additional funding categorically follows the disadvantaged pupils, by providing supplementary services to them in ‘*targeted assistance*’ programmes.

A separate categorical provision aiming to assist foreign/repatriated pupils with limited competence in the Greek language is essential to make them fully comprehend the mainstream curriculum and to alleviate the negative effect of ethnicity on their educational attainment, and for some ethnic groups and subjects on their progress. These differences should be expected as they are related to the fact that foreign/repatriated pupils join the Greek school with poor language skills in Greek (their second language).

A categorical support-teaching provision in mathematics for both Greek and foreign/repatriated pupils with low attainment but who are competent in the Greek language can be justified on the basis of equal educational opportunities, since such a provision would alleviate the effect of social class.

Several studies carried out in other countries demonstrated that initial attainment differences between disadvantaged and advantaged groups increase over time and hence the provision of support teaching is crucial in alleviating the effects of ethnicity and social class.

The National Child Development Study (Davie et al., 1972) found significant class differences in attainment at the age of seven. Children from the unskilled manual social class category (five) were found to have five times more reading problems than children in category one, representing professional and managerial professions, with the difference doubling by age 11. In England, Strand (1999) also found that differences in pupils' educational attainment related to ethnic group, economic disadvantage and sex occur as early as 7 years of age. In Strand's (*op. cit.*) study sex and class differences increased over time since negative progress rates were identified at the end of Key Stage

One. In Scotland, Cowie and Croxford (1999) identified inequalities in pupils' baseline attainment at the beginning of the 1st grade of primary school by entitlement to free school meals (FSM), which is used in Britain as a proxy variable for social class. In addition, the authors found that pupils with free school meal entitlement (FSM) made less progress in reading than other pupils in the 1st year of primary school. Furthermore, pupils' *'initial attainment scores are the most powerful predictors of attainment scores at the end of a period of schooling'* (Goldstein, 1997, p. 382).

Since the current study also confirmed the importance of initial attainment score as the most significant predictor of pupils' final attainment score and demonstrated under-achievement patterns in terms of initial attainment, final attainment and lower progress rates for disadvantaged groups in relation to their more advantaged peers in certain subjects, if disadvantaged pupils with initially significantly lower attainment are not assisted in catching up with the majority group, they may make less progress in relation to advantaged pupils. Negative progress rates entail that in the long run their attainment gap with the more advantaged pupils gets wider. This statement holds for pupils from lower social class in mathematics, for Albanian pupils as well as for pupils with 'other' ethnic background in mathematics, for pupils originating from the democracies of the former USSR and for boys in language. For these groups attainment differences with their more advantaged peers are expected to increase, unless schools take steps to counteract the effects of disadvantage. Additional support teaching to cater for the learning gaps of pupils from disadvantaged backgrounds can be perceived as a vital component in any equal educational opportunities programme, aiming to alleviate the effects of educational and social disadvantage.

Educational leadership plays an important role in Greek support settings, because principals, after consulting with their colleagues, apply early on for the creation of remedial and support schemes in the school. According to Scheerens and Bosker (1997, p. 107) principals should be able to support innovation strategies and stimulate effective instruction.

On the other hand Greek districts play a crucial role, as they allocate funds and personnel to the school according to principals' applications. Often due to personnel shortages, lack of funding and inadequate monitoring, districts fail to appoint teachers to the schools in need, even after having received a favourable application from the principal or districts fail to appoint supply teachers and as a result support class teachers often replace mainstream teachers on leave.

A comprehensive framework for intervention should mean that all pupils' needs should be catered for, in one way or another, through *a diversified approach* that would take into account the different support needs in place.

Such a separate treatment can be provided if foreign/repatriated pupils with limited competence in Greek are placed in categorical support schemes for them such as coach classes and reception classes; on the other hand low achieving foreign/repatriated pupils who are competent in Greek could be placed in additional support schemes along with Greek low achievers who in the majority come from lower social class backgrounds. Such support schemes tailored to pupils' needs still have to be legislated for.

In addition, a further act launched by the Greek Ministry of Education could clarify the tasks that can be undertaken within the whole-day school framework. Such new legisla-

tion delivered in a ministerial decree could provide for the operation of coach classes or reception classes, as well as classes catering for the needs of low achieving pupils within the whole-day school framework.

5. 5. The impact of other variables

For both subjects, the remaining variables tested had a significant impact on pupils' final attainment score. The variables '*number of friends*', '*number of absences*', '*mother's educational level*', '*nursery school attendance*' and finally '*hours of homework in a week*' had an impact on pupils' final attainment levels. (The more friends a pupil had, the less absences s/he had during the school year, the higher his/her mother's educational level, the higher his/her final attainment in mathematics and in language). Also a pupil enrolled in nursery school is more likely to attain more highly in both subjects.

5. 6. Significant interactions in attainment and progress models

The interaction terms '*being a foreign/repatriated pupil by the days of absence during the school year*' and '*being a foreign/repatriated pupil by the number of pupils having named him/her to be their friend*' have an additional impact on foreign pupils' final attainment scores. This impact is over and above the main effects of being foreign, having been absent during the school year, or having friends. These statements hold for both subjects (mathematics and language). The more days foreign/ repatriated pupils have been absent from school, the lower their educational attainment is expected to be. The more friends foreign and repatriated pupils have acquired, the more their final attainment scores are expected to increase. This increase is quicker than the expected increase in final attainment scores of Greek pupils for every additional friend they have.

In terms of progress for mathematics there was a marginally significant positive interaction term between boys and initial score ($p=0.051$), and between minority status (being foreign/repatriated) and initial score. The higher boys' initial attainment scores, the more progress they are expected to make during the school year. The higher the initial scores of foreign/repatriated pupils, the more progress they are expected to make during the school year, in other words the more they will be able to catch up in relation to the majority group. Boys who have high initial attainment (at the beginning of the school year) are expected to make more annual progress than boys who were low achievers at the beginning.

In terms of progress for language, the only significant interaction term identified was '*pupils from lower social class by their initial attainment score*'. Pupils from lower social class who are high achievers at the beginning of the school year are expected to make more progress during the school year than pupils from lower social class who are low achievers.

5. 7. Implications for further study

Further longitudinal school effectiveness research, which would employ a larger sample of primary schools drawn from a variety of areas in Greece (rural and urban) is needed, in order to show whether school and classroom effects exist and in order to verify the extent to which the current findings concerning differences in final attainment and in progress between pupil groups can be identified on a wider basis in basic skills. Such a study should be replicated from time to time, in order to establish whether under-achievement patterns shrink or widen as years go by, and be related to policies regulating the provision of equal educational opportunities to disadvantaged groups of pupils. Attainment and progress differences may also change over time as a result of the waves of immigration of different ethnic groups (e.g. Kurds, Pakistanis) into Greece.

Furthermore it could be examined whether attainment and progress differences between pupil groups can also be identified for other subjects, as for example for science.

Given that the current study has shown that sometimes in language and sometimes in mathematics certain pupil groups made less progress than others, it can be expected that attainment differences between these pupil groups and their respective comparison groups would increase in secondary school. In addition, it would be useful to establish whether the negative impact of belonging to these groups on pupils' progress persists through subsequent years and what that implies for attainment differences for these groups. Hence, it would be interesting to establish whether attainment differences between foreign/repatriated and Greek pupils, between pupils from lower and pupils from higher social class and between boys and girls have increased at the end of year 9,

which marks the end of compulsory education in Greece, or at the end of year 12, which marks the end of secondary education. Such a study may identify different attainment and progress differences between foreign ethnic groups and the majority group from the ones currently identified.

According to Sammons et al. (1995b) previous institutional membership exerts a long-term influence on pupils' attainment and on their progress rates during secondary school. Sammons et al. (*op. cit.*) stated that previous institutional membership was at the origin of differences in attainment and in progress during secondary schooling. In the UK effects associated with pupils' enrolment in a given primary school may be detectable throughout secondary school. Given that in the Greek case classroom effects instead of school effects are identified, if Sammons et al.'s findings are translated into the Greek context, they may imply that classrooms exert a long-term influence on pupils' attainment and progress, even over secondary school years.

Provided that classroom effects can also be identified in future studies, the relationship between instructional correlates, classroom climate and classroom effectiveness for the Greek setting can be explored.

If the Greek government aims to employ a monitoring mechanism that would regularly assess the effectiveness of schools and classrooms, regular school effectiveness exercises have to be carried out at key points in pupils' schooling. Such key points may be the end of each educational level, as well as some point at the beginning of pupils' schooling that would serve as a baseline measure, thus enabling the estimation of pupils' progress rates. An example of how such a monitoring mechanism operates is pro-

vided by the English system. In England the National Foundation for Educational Research (NFER) collects and analyses data from regular assessment exercises conducted at the beginning of Key Stage 2 and at the end of Key Stages 2, 3, and 4. All English students are assessed in basic skills on national attainment targets at ages 7, 11, 14 and 16 through performance-based (normative) tests. In this manner a closer view of the English educational system is achieved.

In Greece, pupils' attainment may be measured at the beginning and at the end of grades that are considered as entry or transition points in primary and/or secondary schooling. For example, such an exercise may involve the 2nd and 6th grades of primary school, the 3rd grade of secondary education, first cycle (9th grade) and the 3rd grade of secondary education, 2nd cycle (12th grade). However, such a system may prove to be too expensive to be adopted by the Greek government at a national level. Some of the inherent merits that such an exercise would entail are that classrooms/ schools performing at a below-average level in terms of progress could be promptly identified, as well as disadvantaged groups who underachieve in them and/or in every school. Such identification could be coupled with school improvement initiatives, as well as with initiatives aiming to alleviate educational disadvantage, such as the provision of categorical support schemes to disadvantaged groups. Such a monitoring exercise if coupled with intervention programmes would have positive effects in low-performing state primary and secondary schools and classrooms, raising standards and alleviating structural inequalities built into the system. Such an exercise is expected to promote equal educational opportunities for all pupils and especially for disadvantaged sections of the school population.

Given that carrying out annual longitudinal census surveys is an expensive mega-exercise another more feasible option would be for districts or research institutes affiliated to the Greek Ministry of Education to carry out smaller School Effectiveness studies based on one area or district only, with special emphasis on educational priority areas. For example regions, cities, or extended geographical areas may be selected on an educational disadvantage principle; for example urban areas in which a high percentage of foreign/repatriated or lower social class pupils are enrolled or remote areas affected by the relative scarcity of resources and teachers. Examples of such urban areas may be Piraeus (the area this study was based on), Menidi, Thrace (an area in which a high percentage of pupils from the Muslim minority is located), or the Cyclades islands (which are affected by both the scarcity factor and relatively high concentration of minorities). Many other areas where teacher shortages are reported, or where it is known that severe educational problems of other kinds exist could be selected for a school effectiveness exercise.

A sample of schools and classrooms located in educational priority areas could be evaluated in relation to schools located in non-priority areas using a longitudinal school effectiveness design. In cases where the purpose of a SER exercise is not to locate individual classrooms and schools performing at a below-average level but to identify differences between groups, nationally representative samples of schools may be derived through matrix-sampling. Such a sampling technique is employed by the California Assessment programme (HMI, 1992, p. 16) and by the assessment system of New Zealand (Irving, 1992, p. 3). In other words a school effectiveness methodology can be perceived as an evaluative tool, which is in a position to identify where problems exist as

well as structural inequalities built into the system, so that steps can be taken to alleviate them. The Greek Ministry of Education and the key stakeholders (districts, and schools) assisted by University departments can realise, prompt and investigate further in schools and areas where problems exist and accordingly take action to alleviate them. This suggestion accords with the suggestion made by Verdis (2002, p. 324) in his thesis that in Greece there is a *'need to focus on educational problems at the local level'*.

For such school effectiveness exercises to be carried out, normative tests measuring pupils' attainment in the Greek language and in mathematics for certain key grades at the end of which regular assessment of pupils would be carried out, should be available as only normative tests can be employed to undertake valid comparisons between pupil groups. Performance in basic skills may be given priority in relation to other lessons. These normative performance-based tests should be based on the curriculum taught at key grades or transition points in primary and/or lower secondary school. Researchers specialising in the field of psychometrics should be delegated the task of creating performance -- based tests along with other research instruments, or of adapting available instruments used in British, or American literature to conditions associated with the functioning of schools or classrooms in Greece. For example, such instruments may pin down school organisational arrangements, school climate, classroom processes, and classroom climate. These instruments need to be highly sensitive to reflect the specific conditions encountered in the highly centralised Greek setting.

Such an exercise needs to be undertaken by a research team, working for a research institute or for the Greek Ministry of Education. The differentiated roles assigned to the team members can allow the school effectiveness study to encompass multiple layers,

instruments and variables. Then aspects related to policy, school and classroom processes and school organisation can be described through these instruments and their impact on pupils' attainment and progress can be estimated.

School Effectiveness studies could investigate and describe the relationship between the progress rates of classrooms and/or schools and the educational processes going on in these settings, during the period between the two measurement occasions.

Other School Effectiveness studies could also employ case studies in order to focus on classrooms/schools identified as positive or negative outliers (performing at an above-average or below-average level) according to a progress model. Such case studies could look into these classroom/school processes in more depth. *'Case studies are sources of critical reflection to be adapted and used by teachers in other schools in ways appropriate to their contexts'* (Lauder et al., 1998, p. 65).

Furthermore, a school effectiveness study could examine the impact of educational innovations on pupils' educational outcomes. Such a school effectiveness study would examine the relationship between the provision of the support instruction in the school, the degree of implementation of these support classes, the curriculum covered there, and foreign/repatriated pupils' educational outcomes in relation to the educational outcomes of majority pupils. Other future studies could assess the impact of schools adopting experimental support teaching programmes to pupils' educational outcomes.

In order to allow for the impact of educational innovations to be validly assessed, these innovations should be implemented in a concise, systematic and uniform way in the schools investigated. *'An important requirement for either comprehensive or narrower*

experimental improvement projects is a rigorous and closely monitored implementation of the treatment' (Scheerens and Bosker, 1997, p. 317).

Such an orientation of SER would inform principals and support class teachers and educational administrators of the impact of educational innovations. For example it can inform them of the impact of coach classes and reception classes on foreign/repatriated pupils' progress in basic skills. Given that organisational arrangements found in coach classes and in reception classes vary, knowing which one maximises pupils' educational outcomes may lead to a more profitable employment of the resources available. Also knowing that the adoption of certain experimental support programmes lead to an increase in pupils' outcomes or knowing which organisational arrangements in experimental support schemes can be at the origin of greater progress is valuable, as such a knowledge can lead policy makers and academics to modify the arrangements in the existing support schemes and design new support schemes that would raise the educational outcomes of pupil groups defined by ethnicity, social class and gender.

REFERENCES

- Afshartous, D. (1995). "Determination of sample size for multilevel model design." Paper presented in the AERA meeting, San Francisco, CA, April 1995.
- Aitkin, M. and R. Zuzovsky (1994). "Multilevel Interaction Models and their Use in the Analysis of Large-scale School Effectiveness Studies." *School Effectiveness and School Improvement* 5(1): 45-73.
- Alexander, K., Fennessey J., McDill, E., D' Amico , R. (1979). School SES influences: Composition or Context? *Sociology of Education* 52(4), Oct 1979: 222-237
- Alexander, K. and B. Eckland (1980). The 'Explorations in Equality of Opportunity' Survey of 1955 High School Sophomores. *Research in Sociology of Education and Socialisation*. A. C. Kerckhoff. Greenwich CT, JAI Press. I, *Longitudinal Perspectives on Educational Attainment*: 31-58.
- Alexander, R. (1996). *Other Primary Schools and Ours: Hazards of International Comparison*. Warwick, UK, University of Warwick, UK: 33.
- Allen, V. G. (1991). Teaching Bilingual and ESL Children. *Handbook of Research on Teaching the English Language Arts*. J. M. J. James Flood, Diane Lapp & James R. Squire. New-York, Macmillan.
- Altrichter, H. and S. Salzgeber (2000). *Some elements of micro-political theory of school development. Images of Educational Change*. H. Altrichter and J. Elliot. Buckingham, Open University Press: 99-110.
- American Department of Education (1998). *Turning Around Low-Performing Schools: A Guide for State and Local Leaders*, American Department of Education.
- American Federation of Teachers (2002). *Questions and Answers about student Testing/Accountability*. <http://www.aft.org/esea/downloads/QAstudenttestingacct.pdf>, American Federation of Teachers.
- Anastasi, A. (1988). *Psychological Testing*. New York, Macmillan Publishing Company.
- Anderson, C. (1982). "The Search for School Climate: A review of the Research." *Review of Educational Research* 52(3): 368-420.
- Anthoulas, T. (1991). Four tools for a contemporary education (in Greek). *Eleutherotypia*. Athens: 35.
- Arnot, M., J. Gray, et al. (1998). *Recent Research on Gender and Educational Performance*. London, The Stationary Office.
- Askew and D. Wiliam (1995). *Recent Research in Mathematics Education 5-16*. London, OFSTED.

- Askew, M. (2001). *What does it mean to learn? What is effective teaching? 'Principles and Practices of Arithmetic Teaching' Innovative approaches for primary classroom*. J. Anghileri. Buckingham - Philadelphia, Open University Press: 134146.
- Baker, C. (1996). *Foundations of Bilingual Education and Bilingualism*. Clevedon, Multilingual Matters.
- Banks, J. A. (1996). *Multicultural education, transformative knowledge, and action: historical and contemporary perspectives*. London, Teachers College.
- Bartholomew, D. J. and M. Knott (1999). *Latent Variable Models and Factor Analysis*. London, Arnold.
- Beath, J. M. (1999). *Schools Must Speak for Themselves: The Case for School Self-Evaluation*. London, Routledge.
- Bennett, E. D. N. (1990). *Talking and learning in groups*, Macmillan, Education.
- Bennett, N. and A. Harris (2001). School Effectiveness and School Improvement: Future Challenges and Possibilities. *School effectiveness and school improvement: alternative perspectives*. A. Harris and N. Bennett. London, Continuum: 177-183.
- Berdie, D. R. and J. F. Anderson (1974). *Questionnaires: Design and Use*. Minnesota, The Scarecrow Press, Inc. Metuchen, N. J.
- Bernstein, B. (1972). A socio-linguistics approach to socialization: With some reference to educability. *Directions in sociolinguistics*. J. J. Gumperz and D. Hymes. New York, Holt, Rinehart and Winston, Inc.
- Blank, R. K. and J. G. Manise (2000). "State Education Indicators for Monitoring Progress of Title I: A New Annual Series of Reports." *Journal of Education for Students Placed at Risk* 5(3): 215-223.
- Blatchford, P. and P. Mortimore (1994). "The issue of class size in schools: what can we learn from research?" *Oxford Review of Education* 20(4): 411-428.
- Bloom, B. S., J. T. Hastings, et al. (1971). *Handbook on Formative and Summative Evaluation of Pupil Learning*. New York, McGraw-Hill.
- Boaler, J. (1997). "Setting, Social Class and Survival of the Quickest." *British Educational Research Journal* 23(5): 575-595.
- Bochner, S. (1996). "The learning strategies of bilingual versus monolingual students." *British Journal of Educational Psychology* 66: 83-93.
- Bollen, R. (1996). School effectiveness and school improvement: The intellectual and policy context. *Linking school effectiveness and school improvement*. D. Reynolds, R. Bollen, B. Creemers et al. London, Routledge: 1-20.

- Bondi, L. (1991). "Attainment at Primary Schools: an analysis of variations between schools." *British Educational Research Journal* 17(3): 201-217.
- Bone, A. (1983). *Girls and Girls Only Schools: A Review of the Evidence*. Manchester; Equal Opportunities Commission.
- Borman, G. D. and L. T. Rachuba (2001). *Academic success among poor and minority students: An analysis of competing models of school effects*. Baltimore, MD:, Johns Hopkins University, Center for Research on the Education of Students Placed At Risk.
- Borman, G. D., K. K. Wong, et al. (2001). Coordinating Categorical and Regular Programs: Effects on Title I Students' Educational Opportunities and Outcomes. *Title I: Compensatory Education at the Crossroads*. S. C. S. Geoffrey, D. Borman, Robert E. Slavin: Editors. Mahwah, New Jersey, London, Laurence Erlbaum Associates: 79-116.
- Bosker and Scheerens (1989). "Issues in the interpretation of the effects of school effectiveness research." *International Journal of Educational Research* 13: 741-751.
- Bosker, R. J. (1990). "School and Instruction Effects on Mathematics Achievement." *School Effectiveness and School Improvement* 1(4): 233-248.
- Bottery, M. (2001). School Effectiveness, School Improvement and the Teaching Profession of the Twenty-first Century. *School effectiveness and school improvement: alternative perspectives*. A. Harris and N. Bennett. London, Continuum: 140-160.
- Bray, M. and M. Thomas (1995). "Levels of comparison in Educational Studies: Different Insights from Different Literatures and the Value of Multilevel Analyses." *Harvard Educational Review* 65(3): 472-490.
- Brandsma, H. and Knuver, J. (1989). "Effects of school and classroom characteristics on pupil progress in language and arithmetic." *International Journal of Educational Research* 13: 777-788.
- Brandsma, H., and Doolaard (1999). "Differences in effectiveness between primary schools and their impact on secondary school". *School effectiveness and School Improvement*, 10: 430-450.
- Breen, R. and D. B. Rottman (1995). *Class Stratification: A Comparative Perspective*. London, Harvester Wheatsheaf.
- Brighouse, T. and D. Woods (1999). *How to improve your school*. London, Routledge.
- Brimer, A., G. Madaus, et al. (1978). *Sources of Differences in School Achievement*. London, NFER Publishing Company.
- Brookover, W. B., Schweitzer, J. H., Schneider, J. M., Beady, C. H., Flood, P. K., and Wisenbaker, J. M. (1978). Elementary school social climate and school achievement. *American Educational Research Journal*. 15, 301-318.

- Brookover, W. B., C. Bready, et al. (1979). *School social systems and student achievement - schools can make a difference*. New York, Praeger Publishers.
- Brookover, W. (1985). "Can we make schools effective for minority students?" *Journal of Negro Education* 54(3): 257-68.
- Brophy, J. and T. Good (1986). Teacher behavior and student achievement. *Handbook of research on teaching*. M. Wittrock, Macmillan: 340-370.
- Brophy, J. (1987). *Research on teacher effects: Uses and Abuses*. Michigan, The Institute for Research on Teaching, College of Education, Michigan State University, East Lansing: 43.
- Brown, S., S. Riddell, et al. (1996). Possibilities and problems of small-scale studies to unpack the findings of large - scale studies of school effectiveness. *Merging traditions: the future of research on school effectiveness and school improvement*. J. Gray, D. Reynolds, C. Fitz-Gibbon and D. Jesson. London, Cassell: 93-120.
- Bryk, A. and S. Raudenbush (1992). *Hierarchical Linear Models: Applications and Data Analysis Methods*. Newbury Park, CA, Sage.
- Bushman, J., G. Goodman, et al. (2001). "California Testing: How Principals Choose Priorities." *Educational Leadership* 59(1): 33-36.
- California Department of Education (1999). *Designing a Standards-Based Accountability System for Language Minority and Immigrant Student Populations*. California, California Department of Education.
- California Department of Education (1999). *Title I Program Improvement (PI) in California: Questions and Answers*. California, California Department of Education: 1-7.
- California Department of Education (2002). *Low-Performing Schools*. 2002.
- California Department of Education (2002). *Title I Program Improvement in California: Frequently Asked Questions*. 2002.
- California Department of Education (2002). *Standards-based Reform in California*. California, California Department of Education: 1-4.
- California Department of Education Policy and Evaluation Division (2002). *Academic Performance Index*, California Dept of Education, Policy and Evaluation Division. 2002.
- Camilleri A. (2003). The implementation of poverty and social exclusion indicators for Malta: a case study. *Social Statistics in the Enlarged EU: 89th DGINS Conference*, Athens, European Communities, Eurostat. Luxemburg: 35-42.
- Carpenter, P., Hayden, M. (1987). "Girls' Academic Achievements: Single-Sex Versus Coeducational Schools in Australia." *Sociology of Education* 60(3): 156-167.

- Carpenter, T. P., E. Fennema, et al. (1989). "Using Knowledge of Children's Mathematics Thinking in Classroom Teaching: An Experimental Study." *American Educational Research Journal* 26(4): 499-531.
- Carpenter, T. P., E. Fennema, et al. (2000). *Cognitively Guided Instruction: A research-Based Teacher Professional Development Program for Elementary School Mathematics*. Madison, NCISLA Wisconsin Center for Education Research: 9.
- Carroll, J. (1963). "A model of school learning." *Teachers College Record* 64: 723-733.
- Carroll, B. (1989). "The Carroll model, a 25-year retrospective and prospective view." *Educational Researcher* 18: 26-31.
- Cawelti, G. (1995). *Handbook of Research on Improving Student Achievement*. Arlington, Educational Research Service.
- Central Advisory Council for Education (England) (1967). *Children in their primary schools. The Plowden report, Volume 1: Report*, London, Crown copyright
- Chall, J. S. and M. E. Curtis (1991). *Children at risk. Handbook of Research on Teaching the English Language Arts*. J. Flood, J. M. Jensen, D. Lapp and J. R. Squire. New York, Macmillan: 349-371.
- Chall, J. S. and M. E. Curtis (1991). Responding to individual differences among language learners. *Handbook of research on teaching the English language arts*. J. Flood, J. M. Jensen, D. Lapp and J. Squire. Toronto, Collier MacMillan: 349-355.
- Chapman, J. D. (1991). *The effectiveness of schooling and of educational resource management. Monograph no.1, A conceptual-analytical framework*. Paris, OCDE: 155.
- Chitty, C. (1997). "The school effectiveness movement: origins, shortcomings and future possibilities." *The curriculum journal* 8(1): 45-62.
- Chrispeels, J. and S. Pollack (1989). Equity in schools and Equity districts. '*School Effectiveness and School Improvement*'. B. Creemers, T. Peters and D. Reynolds. Amsterdam / Lisse, Swets & Zeitlinger.
- Cohen, L. (1976). *Educational research in classrooms and schools*.
- Coleman, J. S., Campell, E. (1966). *Equality of Educational Opportunity*. Washington:, US Government Printing Office.
- Coleman, P. K., Byrd C. P. (2003). Interpersonal correlates of peer victimisation among young adolescents, *Journal of youth and adolescence*, 32(4), 301-314.
- Coley, R. (2002). *An uneven start: indicators of inequality in school readiness*. Princeton, Education Testing Service, Statistics and Research Division, Policy Information Report: 64.
- Comer, J. P. (1989). "Parent participation: Fad or function?." *Educational Horizons* 69: 182-188.

- Commission for Racial Equality (1976). *Educational Needs of Children from Minority Groups*: 21.
- Connecticut State Board of Education, Office of Board Matters (2004), State Board of Education Goals and Comprehensive Plan for Education, <http://www.state.ct.us/sde/board>
- Cooper, H. (1989). "Synthesis of research on Homework." *Educational Leadership* (47) November 1989: 85-91.
- Cotton, K. (1995). *Effective schooling practices: A research synthesis*. Northwest. Regional Education Laboratory.
- Couloubaritsis, A. and P. Kavouri (1994). "The curriculum implementation of history in primary schools (in Greek)." *Contemporary Education (Greek journal)*(74): 62-70.
- Council of Local Education Authorities (CLEA) (1997). *Local Education Authorities and Improving Schools*. London, The Council of Local Education Authorities.
- Cowie, M. and Croxford, L. (1999). *Using Baseline Assessment for Early Intervention*. Centre for Educational Sociology, University of Edinburgh.
- Creemers, B., T. Peters, et al. (1989). *School Effectiveness and School Improvement*. Amsterdam/Lisse, Swets & Zeitlinger.
- Creemers, B. (1992). *School effectiveness and effective instruction: The need for a further relationship*. In Bashi J. and Sass Z. 'School Effectiveness and Improvement: Selected proceedings of the Third International Congress for School Effectiveness': Jerusalem, Israel: Magnet Press. 105-132
- Creemers, B. (1994a). The History, value and purpose of school effectiveness studies. In Reynolds, D., Creemers, B., Nesselrodt, P. Schaffer, E., Stringfield, E., and Teddlie, C. 'Advances in School Effectiveness Research and Practice'. Oxford, Pergamon.
- Creemers, B. (1994b). *The effective classroom*. London, Cassell.
- Creemers, B. (1996). *The goals of school effectiveness and school improvement*. in D. Reynolds, R. Bollen, B. Creemers et al. 'Making good schools: Linking school effectiveness and school improvement'. London, Routledge: 21-35.
- Creemers, B. (1996). *The school effectiveness knowledge base* in D. Reynolds, R. Bollen, B. Creemers et al., *Making good schools: Linking school effectiveness and school improvement*'. London, Routledge: 36-58.
- Creemers, B. and G. Reezigt (1997). "School Effectiveness and School Improvement. Sustaining Links." *School Effectiveness and School Improvement* 8(4): 396-429.
- Creemers, B., S. Stringfield, et al. (2002). The quantitative data. *World class schools: International perspectives on school effectiveness*. D. Reynolds, B. Creemers, S. Stringfield, C. Teddlie and G. Schaffer. London, Routledge / Falmer: 33-55.

- Creswell, J. W. (1994). *Research Design: qualitative and quantitative approaches*. London, Sage.
- Creswell, J. (1998). *Qualitative inquiry and research design: choosing among five traditions*. Thousand Oaks, Calif, Sage Publications.
- Croll, P. (1996). Teacher-Pupil Interaction in the Classroom. *Effective Primary Teaching*. P. Croll and N. Hastings. London, David Fulton Publishers: 14-28.
- Croll, P. and N. Hastings (1996). *Effective primary teaching: research-based classroom strategies*. London, David Fulton Publishers Ltd.
- P. Croll and N. Hastings (1997). Teachers Matter. *Effective Primary Teaching*. P. Croll and N. Hastings. London, David Fulton Publishers: 1-13.
- Croll, P. (1997). Teaching and Research. *Effective Primary Teaching*. P. Croll and N. Hastings. London, David Fulton Publishers: 133-138.
- Cronbach, L., R. Linn, et al. (1997). "Generalisability analysis for performance assessments of student achievement or school effectiveness." *Educational and Psychological Measurement* 57(3, June): 373-399.
- Croydon (1994). *Raising achievement in mathematics at Key Stage 2*, Education Department, England: 9.
- Cummins, J. (1984). Wanted: A theoretical framework for relating language proficiency and academic achievement among bilingual students. *Language Proficiency and Academic Achievement*. C. Rivena. Clevedon, Avon, Multilingual Matters, LTD: 2-19.
- Cummins, J. (1984). *Bilingualism and special education: Issues in assessment and pedagogy*. Clevedon, England, Multilingual Matters Ltd.
- Cummins, J. (1984). Underachievement among ethnic minority children. *Bilingualism and special education. Issues in assessment and Pedagogy*. J. Cummins. London, Multilingual Matters Ltd.
- Cuttance, P. (1994). "Monitoring Educational Quality through Performance Indicators for School Practice." *School Effectiveness and School Improvement* 5(2): 101-126.
- Dale, R. (1997). The State and the Governance of Education: an analysis of the restructuring of the state-education relationship. *Education: Culture, Economy, Society*. A. Halsey. Oxford: 273-282.
- Damanakis, M. (2001). *Education of repatriated and foreign pupils in Greece. An Intercultural approach* (in Greek). Athens, Gutenberg.
- Damme, J. V., P. Onghena, et al. (2001). *Effects of schools and classes upon mathematics, language achievement and non-cognitive outcomes*. The 14th International Congress for School Effectiveness and Improvement, Toronto, Canada, http://www.edu.yorku.ca/ICSEI/papers/Toronto_text.html.

- Deem, R. (1986), Gender and social class in Rogers, R., *Education and social class*, The Falmer Press, London: 173-191
- Denvir, B., C. Stolz, et al. (1982). *Low attainers in mathematics 5-16: policies and practices in schools*. London, Methuen Educational.
- Department for Education and Employment (DfEE) 0253/2000 (2000). *Recognising Progress; Getting the most from your data*, Department for Education and Employment. 2000.
- Department for Education and Science (DfES) (2002). *2002 KS1-KS2 Value-Added Pilot*, Department for Education and Science (DfES).
- Department for Education and Employment (DfEE) (1997). *Excellence in Schools*. London, Department for Education and Employment: 84.
- Department for Education and Employment (DfEE) (2002). *Let's make learning cool*, DfEE. 2002.
- Department for Education and Employment (DfEE) (2002). *The Government response to boys' underachievement*. <http://www.literacytrust.org.uk/Database/boys/boysgovt.html>, DfEE.
- Department for Education and Employment (DfEE) (2002). *Let's make learning cool: Classroom issues*. http://www.dfee.gov.uk/teacher/data/issues/data/cool_10.htm, DfEE.
- Department for Education and Employment (DfEE) (2002). *School Improvement Grant 2000-2001: Q&A*
What can the School Improvement Grant be used for? DfEE. 2002.
- Derrington, C. (2000). *The LEA contribution to school improvement - a role worth fighting for*. Slough, National Foundation for Educational Research: 124.
- Dixon-Kraus, L. (1996). *Vygotsky in the classroom: Mediated literacy instruction and assessment*. White Plains, NY, Longman.
- Doyle, W. (1986). Curriculum organisation and management. *Handbook of research on teaching*. Wittrock. New York, Macmillan.
- Duncan, C. (1994). The Multicultural Dimension of Quality and Equality. *Improving Education: Promoting Quality in Schools*. P. Ribbins and E. Burrige. London, Cassell: 105-113.
- Edmonds, R. (1979). "Some Schools Work and More Can." *Social Policy* 9: 28-32.
- Edmonds, R. (1982). "Programs of school improvement: An overview." *Educational Leadership* 40(3): 4-11.
- Einsiedler, W. and G. Treinies (1997). "Effects of Teaching Methods, Class Effects, and Patterns of Cognitive Teacher-Pupil Interactions in an Experimental Study in Primary School Classes." *School Effectiveness and School Improvement* 8(3): 327-353.

- Elliot, J. (1996). "School Effectiveness Research and its critics: alternative visions of schooling." *Cambridge Journal of Education* 26: 199-223.
- Elliot, J. (2000). Towards a synoptic vision of educational change in advanced industrial societies. *Images of Educational Change*. H. Altrichter and J. Elliot. Buckingham, Open University Press: 175-220.
- Ellis, R. (1994). *The study of second language acquisition*. Oxford, Oxford University Press.
- Erikson, R. (1984). "Social Class of Men, Women and Families." *Sociology* 18.
- Farrow, S., P. Tymms, et al. (1999). "Homework and attainment in primary schools." *British Educational Research Journal* 25(3): 323-341.
- Fidler, B. (2001). A Structural critique of School Effectiveness and School Improvement. *School effectiveness and school improvement: alternative perspectives*. A. Harris and N. Bennett. London, Continuum: 47-74.
- Field, T., Diego, M. and Sanders (2001). Adolescent depression and risk factors. *Adolescence*, 36(143): 491-498.
- Firestone, W. A. (1991). Introduction. In Bliss, J. R., Firestone, W. A., and Richards C. E. 'Re-thinking Effective Schools: Research and Practice'. Englewood Cliffs. New Jersey: Prentice Hall.
- Firnberg, D. (1997). *The Cassell spelling dictionary*. London :, Cassell.
- Fitz-Gibbon, C. (1994). Performance Indicators, Value Added and Quality Assurance. *Improving Education*. P. Ribbins and E. Burrige. London, Cassell. 141-149.
- Fitz-Gibbon, C. (1996). Monitoring school effectiveness: simplicity and complexity. *Merging traditions: The future of research on school effectiveness and school improvement*. J. Gray, D. Reynolds, C. Fitz-Gibbon and D. Jesson. London, Cassell: 74-90.
- Fitz-Gibbon, C. and S. Kochan (2001). School Effectiveness and Education Indicators. *The International Handbook of School Effectiveness Research*. C. T. a. D. Reynolds. London and New York, Falmer Press: 257-282.
- Floyd, D. G. (1999). *An Examination of the links between the pedagogical culture of primary schools, school effectiveness and school improvement*. Institute of Education. London, University of London: 366.
- Fraser, B. J. (1989). "Twenty years of classroom climate work: progress and prospect." *Journal of Curriculum studies* 21(4): 307-327.
- Fraser, B. J. (1991). Two decades of classroom environment research. *Educational environments: Evaluation, antecedents and consequences*. B. J. Fraser and H. J. Walberg. Oxford, Pergamon Press.

- Frobisher, L. (1999). *Learning to teach number: a handbook for students and teachers in the primary school*. Cheltenham, Stanley Thornes.
- Gamarnikow, E. and A. Green (1999). Developing Social Capital: Dilemmas, Possibilities and Limitations in Education. Tackling Disaffection and Social Exclusion. A. Hayton. London, Kogan Page: 46-64.
- Gammon, A. (1992). *Bilingual Students Learning Mathematics*. Institute of Education. London, University of London: 109.
- Gandal, M. and J. Vranek (2001). "Standards: Here Today, Here Tomorrow." *Educational Leadership* 59(1): 7-13.
- Ganzeboom, H. and D. J. Treiman (1996). "Internationally Comparable Measures of Occupational Status for the 1998 International Standard Classification of Occupations." *Social Science Research* 25: 201-239.
- Gardner, H. (1983). *Frames of mind: the theory of multiple intelligences*, Basic Books.
- Geert W. J. M. Driessen and Lia W. J. Mulder (1999) 'The enhancement of educational opportunities of disadvantaged children' in '*Enhancing Educational Excellence, Equity and Opportunity*' Roel J. Bosker, Bert P.M. Creemers and Sam Stringfield, ICO, Kluwer Academic Publishers: 37-63
- Gerber, S. B. and J. D. Finn (2001). "Teacher Aides and Students' Academic Achievement." *Educational Evaluation and Policy Analysis* 23(2): 123-143.
- Gibson, A. and S. Asthana (1998). "Schools, pupils and examination results: contextualising school performance." *British Educational Research Journal* 24: 269-282.
- Gillborn, D. and C. Gipps (1996). *Recent research on the achievements of ethnic minority pupils*. London, Office for Standards in Education: 91.
- Gillborn, D. (1999). Fifty years of failure: 'Race' and Education Policy in Britain. *Tackling Disaffection and Social Exclusion*. A. Hayton. London, Kogan Page: 135-155.
- Gillborn, D. and H. S. Mirza (2000). *Educational Inequality: Mapping Race, Class and Gender*. London, OFSTED: www.ofsted.gov.uk: 31.
- Gillborn, D. and D. Youdell (2000). *Rationing Education*. Buckingham, Philadelphia, Open University Press.
- Glenn, J. A. (1977). *Teaching Primary Mathematics*. London, Harper & Row.
- Goldstein, H. (1987). *Multilevel Models in Educational and Social Research*. Oxford, Oxford University Press.
- Goldstein, H., J. Rasbash, et al. (1992). Multilevel models for comparing schools. *Multilevel Modelling Newsletter*. London, Department of Mathematics, Statistics and Computing, Institute of Education, University of London. 4: 5-6.

- Goldstein, H., J. Rasbash, et al. (1993). "A Multilevel Analysis of school examination results." *Oxford Review of Education* 19(4): 425-433.
- Goldstein, H. (1994). Educational quality and student achievement. *Improving Education*. P. Ribbins and E. Burridge. London, Cassell: 150-157.
- Goldstein, H. (1995). *'Interpreting international comparisons of student achievement'*, UNESCO Publishing.
- Goldstein, H. (1995). *Multilevel Statistical Methods*. London, Arnold.
- Goldstein, H. and P. Sammons (1995). *The issue of secondary and junior schools on sixteen year examination performance: A cross-classified multi-level analysis*. London, Institute of Education.
- Goldstein, H. and D. Spiegelhalter (1996). "League tables and their limitations: statistical issues in comparisons of institutional performance." *Journal of the Royal Statistical Society A*(159): 385-443.
- Goldstein, H. (1997). "Methods in School Effectiveness Research." *School Effectiveness and School Improvement* 8(4): 369-395.
- Goldstein, H. (1997). "Assessing the performance of schools: limits and league tables." *Managing Schools today: Value-added tables: the less-than-holly grail*.(6): 18-19.
- Goldstein, H. (1998). *Targets, performance indicators and proper uses of educational data*. London, Institute of Education: 4.
- Goldstein, H. and P. Blatchford (1998). "Class Size and Educational Achievement: a review of methodology with particular reference to study design." *British Educational Research Journal* 24(3): 255-268.
- Goldstein, H. and G. Woodhouse (2000). "School effectiveness research and educational policy." *Oxford Review of Education* (26): 353-363.
- Goldstein, H., P. Huiqi, et al. (2000). *The use of value added information in judging school performance*, Hampshire Local Education Authority. 2000.
- Goldstein, H. and A. Heath (2000). *Educational Standards*. Oxford, Oxford University Press.
- Goldstein, H. (2001). "Using Pupil Performance Data for Judging Schools and Teachers: scope and limitations." *British Educational Research Journal* 27(4): 433-442.
- Goldstein, H. (2001). "League Tables and Schooling." *Science in Parliament, Vol. 58* No.2 Whit. 2001, pp.4-5 58(2): 4-5.
- Goldstein, H., W. Browne, et al. (2001). "Extensions of the Intra-unit correlation coefficient to complex generalised linear multilevel models." multilevel modelling webpage: 12.
- Goldstein, H. and P. Noden (2001). *Modelling Social Segregation*. Educational Choice Conference, Centre for the Economics of Education.

- Goldstein, H., W. Browne, et al. (2002). *Partitioning variation in multilevel models*. 2002.
- Goldstein, H. (2003). *Multilevel Statistical Methods*. London, Hodder Arnold.
- Good, T. L. and J. E. Brophy (1986). School Effects. *Handbook of Research on Teaching*. M. C. Wittrock. London, Macmillan Publishing Company. 1: 570-601.
- Goodland, J. (1984). *A Place Called School*. New York, McGraw Hill, Inc.
- Goodwin, J. (1995). *Socio-academic factors/ type of school and educational attainment*. University of Kent. Canterbury, of Kent: 97.
- Gotovos, T. (1986). *The logic of the existent school (in Greek)*. Athens, Gutenberg.
- Gravemeijer, K. (2001). *Fostering a dialectic relation between theory and practice. 'Principles and Practices of Arithmetic Teaching' Innovative approaches for primary classroom*. J. Anghileri. Buckingham - Philadelphia, Open University Press: 147-161.
- Gray, J. and B. Wilcox (1995). *Good School, Bad School*. London, Open University Press.
- Gray, J., D. Jesson, et al. (1996). The challenges of school improvement: Preparing for the long haul. *Merging traditions: The future of research on school effectiveness and school improvement*. J. Gray, D. Reynolds, C. Fitz-Gibbon and D. Jesson. London, Cassell: 165-184.
- Gray, J., D. Goldstein, H., and Jesson, D. (1996). Changes and improvements in school effectiveness: trends over five years. *Research Papers in Education*, 11: 35-51.
- Gray J., Jesson D. and Reynolds D. (1999) 'The challenges of School Improvement: Preparing for the long Haul' in Gray J., Reynolds D., Fitz-Gibbon C. and Jesson D. 'Merging traditions: The future of research on school effectiveness and school improvement', London, Cassell.
- Gray, J., H. Goldstein, et al. (2001). "Predicting the future: the role of past performance in determining trends in institutional effectiveness at A level." *British Educational Research Journal* 27(4): 391-405.
- Greene, J. C., V. J. Caracelli, et al. (1989). "Towards a Conceptual Framework for Mixed-Method Evaluation Designs." *Educational Evaluation and Policy Analysis* 11(3): 255-274.
- Griffin, G. A. (1991). Lessons Learned: Closing the Gap Between Effective Schools Knowledge and the Education of Teachers. *School effectiveness and Improvement*. J. Bashi and Z. Sass. Jerusalem, The Magness Press, the Hebrew University. 1.
- Grift, W. V. d., T. Houtveen, et al. (1997). "Instructional Climate in Dutch Secondary education." *School Effectiveness and School Improvement* 8(4): 449-462.
- Grift, W. V. d. and J. F. Lam (1998). *Do schools with better educational processes have higher output?* Congress on School Effectiveness and Improvement '98., Manchester.

- Grisay, A. (1996). *Évolution des acquis cognitifs et socio-affectifs des élèves au cours des années de collège*. Liège, Université de Liège.
- Guilborn, D. and C. Gipps (1996). *Recent Research on the Achievement of Ethnic Minority Pupils*. London, Institute of Education, University of London.
- Hair J., Anderson R., Tatham R. and Black W. (1995): *Multivariate data analysis with readings*. London, Prentice - Hall International, INC.
- Hakuta, K., Y. G. Butler, et al. (2000). *How Long Does It Take English Learners to Attain Proficiency?* Stanford, Stanford University: 28.
- Hallam, S. and I. Toutounji (1996). *What do we know about the grouping of pupils by ability?* London, Institute of Education.
- Halliday, J. (1996). "Diversity and Tradition in Educational Practices." *Educational Research and Evaluation* 2(4): 351-365.
- Hallinger, P. and J. Murphy (1986a). "The social context of effective schools." *American Journal of Education* 94(3): 328-355.
- Haque, Z. and J. F. Bell (2001). "Evaluating the Performances of Minority Ethnic Pupils in Secondary Schools." *Oxford Review of Education* 27(3): 357-368.
- Hargreaves, D., D. Hopkins, et al. (1991). *The empowered school: the management and practice of development planning*. London, Cassell.
- Hargreaves, D. H. (2001). "A Capital Theory of School Effectiveness and Improvement." *British Educational Research Journal* 27(4): 487-503.
- Harland, J. (1996). *Evaluation as Realpolitik. Understanding Educational Research*. London, Routledge.
- Harlen, W. (1992). *The teaching of science*. London, David Fulton Publishers.
- Harlen, W. (2001). *The teaching of science in primary schools*. London, David Fulton Publishers.
- Harris, A. and J. Russ (1994). *Pathways to school improvement*. London, Department of Employment: 19.
- Harris, A., I. Jamieson, et al. (1996). *School Effectiveness and School Improvement: A practical Guide*. London, Pitman Publishing.
- Harris, A. (2001). Contemporary perspectives in School Effectiveness and School Improvement. *School effectiveness and school improvement: alternative perspectives*. A. Harris and N. Bennett. London, Continuum: 7-25.
- Hartle, F. (2000). "Classroom climate: a framework for professional development." *Education review* 1(14): 35-40.

- Hartle, F., K. Everall, et al. (2001). *Getting the best out of performance management in your school*. London, Kogan Page.
- Hatcher, R. and S. Thomas (2000). "Richard Hatcher & Sally Thomas on Equity and School Effectiveness Research." *Race, Ethnicity and Education* 3(1): 103-109.
- Hauser, R. M., W. H. Swell, et al. (1976). High School Effects on achievement. *Schooling and Achievement in American Society*. W. H. Swell, R. M. Hauser and D. L. Featherman. New York, Academic Press.
- Havenman, R. and Wolfe B. (1995). The determinants of children's attainments: a review of methods and findings. *Journal of Economic Literature*. 33: 1829-78
- Headington, R. (2000). *Monitoring, assessment and accountability*. London, David Fulton Publishers.
- Hextall, I. and P. Mahony (1998). Effective Teachers for Effective Schools. *School effectiveness for whom?* R. Slee, G. Weiner and S. Tomlinson. London, Falmer Press.
- Hill, P., K. Rowe, et al. (1995). *Factors affecting students' educational progress: Multilevel modelling of educational effectiveness*. 8th International Congress of School Effectiveness and Improvement, Leewarden, The Netherlands.
- Hill, P. and K. Rowe (1996). "Multilevel Modelling in School Effectiveness Research." *School Effectiveness and School Improvement* 7(1): 1-34.
- Hill, P. W. and K. Rowe (1998). "Modelling Student Progress in Studies of Educational Effectiveness." *School Effectiveness and School Improvement* 9(3): 310-333.
- Hill, P. (2001). *Teaching and school effectiveness*. State of Victoria, Australia, State Government Victoria, Department of Education, Employment and Training: 17.
- HMI (Her Majesty's Inspectorate) (1991). *Aspects of Education in the USA*. Edinburgh, The Scottish Office Education Department, Welsh Office Education Department,: 39.
- Hodgson, A. (1999). Analysing Education and Training Policies for Tackling Social Exclusion. *Tackling Disaffection & Social Exclusion*. A. Hayton. London, Kogan Page Limited: 11-32.
- Hofman, A. (1994). "School Effects on Performances of Minority Pupils." *School effectiveness and School Improvement* 5(1): 26-44.
- Hofman, A. (1995). "Cross-level relationships within effective schools." *School Effectiveness and School Improvement* 6(2): 146-174.
- Holmes, M. and P. Croll (1989). "Time spent on homework and academic achievement." *Educational Research* 31(1): 36-45.
- Holt, A. and P. J. Murphy (1993). "School Effectiveness in the Future: the Empowerment Factor." *School Organisation* 13(2): 175-186.

- Hopkins, D. (1993). *A teacher's guide to classroom research*. Buckingham, Open University Press.
- Hopkins, D. (1994). School Improvement in an era of Change. Improving Education: *Promoting Quality in Schools*. P. Ribbins and E. Burridge. London, Cassell: 74-91.
- Hopkins, D. and N. Lagerweij (1996). The school improvement knowledge base. *Making good schools: Linking school effectiveness and school improvement*. D. Reynolds, R. Bollen, B. Creemers et al. London, Routledge: 59-93.
- Hopkins, D., M. West, et al. (1996). *Improving quality education for all*. London, David Fulton Publishers.
- Hopkins, D., M. West, et al. (1997). *Creating the conditions for classroom improvement*. London, David Fulton.
- Hopkins, D. and D. Reynolds (2001). "The Past, Present and Future of School Improvement: towards the Third Age." *British Educational Research Journal* 27(4): 459-475.
- Huberman, A. M. and M. B. Miles (1994). Data Management and Analysis Methods: *Handbook of Qualitative research*. D. Lincoln. London, Sage: 428-444.
- Hutchison, D. (1993). "School Effectiveness Studies Using Administrative Data." *Educational Research* 35(1): 27-47.
- Iliou, M. (1982). "The meaning of equality of opportunity has become almost equated with centralisation, repetition and uniformity in education." *Review of Political Science* (3): 101-119.
- Iliou, M. (1988). *Educational and Social dynamics (in Greek)*. Athens, Poreia.
- Irving, J. (1992). *Educational Assessment & Monitoring in New Zealand*, New Zealand's Ministry of Education: 6.
- James, C. and U. Connolly (2000). *Effective Change in schools*. London, Routledge Falmer.
- Jamieson, I. and F. Wikelly (2001). A Contextual Perspective: Fitting School Round the Needs of Students. *School effectiveness and school improvement: alternative perspectives*. A. Harris and N. Bennett. London, Continuum: 161-176.
- Jenks, C., M. S. Smith, et al. (1972). *Inequality: A reassessment of the effect of family and schooling in America*. New York, Basic Books.
- Jenks, C., S. Bartlett, et al. (1979). *Who gets ahead? The determinants of economic success in America*. New York, Basic Books.
- Jones, K. and K. Bird (2000). "Partnership as Strategy: public-private relations in Education Action Zones." *British Educational Research Journal* 26(4): 491-506.

- Journal of the Government of the Greek Democracy (FEK) (1983). "Presidential Decree 494." *Journal of the Government of the Greek Democracy* - (FEK), 186(1st): 2493-2690.
- Journal of the Government of the Greek Democracy (FEK) (1983). "Law 1404." *Journal of the Government of the Greek Democracy* - (FEK), 173(1st): 1562-2568.
- Journal of the Government of the Greek Democracy (FEK) (2000). "Law 2817." *Journal of the Government of the Greek Democracy* - (FEK), 78(1st): 1565-1602.
- Journal of the Government of the Greek Democracy (FEK) (1990). "Law 1894." (*Journal of the Government of the Greek Democracy* - (FEK)), 110.
- Journal of the Government of the Greek Democracy (FEK) (1997). "Law 2525/1997." *Journal of the Government of the Greek Democracy* - (FEK), 188 (1st issue /188).
- Journal of the Government of the Greek Democracy (FEK) (2000). "Law 2817/2000." *Journal of the Government of the Greek Democracy* - (FEK), 78 (1st issue /78).
- Joyce, B. and B. Showers (1988). *Student Achievement Through Staff Development*. New York, Longman.
- Joyce, B. and N. M. Weil (1996). *Models of Teaching*. Englewood Cliffs, Prentice-Hall.
- Jurs, W. W. S. G. (1985). *Educational Measurement and Testing*. Boston, London, Sydney, Toronto, Allyn and Bacon.
- Kaluge, L. (1998). *Some Factors Related to Educational Attainment in Indonesian Primary Schools*. Institute of Education. London, University of London.
- Kavouri, P. (1996). *Innovations in curriculum, and administration, school climate and other factors in Greek Primary Schools. Their relationship and impact on students' outcomes*. PhD thesis, College of Cardiff. Cardiff, Wales.
- Kerckhoff, A., Mason W and Sandomirsky Poss S (1973). On the accuracy of children's reports of family social status. *Sociology of Education*, <http://uk.jstor.org>: 46, Spring 1973, 219-247.
- King, R. (1983). *The sociology of school organisation*. London, Methuen.
- Kleinbaum, D. G., Kupper L., Muller K. E., and Nizam A. (1998). *Applied Regression Analysis and Other Multivariate Methods*, Duxbury Press, Pacific Grove: 788.
- Koeler, M. S. and D. A. Grows (1992). *Mathematics Teaching Practices and Their Effects. Handbook of Research on Mathematics Teaching and Learning*. D. Grows. New York, MacMillan: 115-126.
- Korilaki, P. (1997). *An Examination of the Performances of Minority-Repatriated Students in Compulsory Education in Greece: an Exploratory Study*. Department of Educational Studies. Oxford, University of Oxford: 105.

- Korilaki, P. (forthcoming). *Describing the operation of coach classes and reception classes*. Σύγχρονη Εκπαίδευση
- Kosmopoulos, A. (1985). "The school we miss (in Greek)." *Educational Review*: 1-41.
- Krashen, S. (1982). *Principles and Practice in Second Language Acquisition*. Oxford, Pergamon Press.
- Krashen, S. (1985). *The Input Hypothesis: Issues and Implications*. New York, Longman.
- Krashen, S. D. and T. D. Terrel (1988). *The natural approach: Language Acquisition in the Classroom*. London, Prentice Hall.
- Kreber, C. and P. Brook (2001). "Impact Evaluation of Educational Development Programmes." *The International Journal for Academic Development*.
- Kretovics, J., K. Farber, et al. (1991). "Reform from the bottom up: Empowering Teachers to Transform Schools." *Phi Delta Kappan*: 295-299.
- Kyriakides, L., R. J. Campbell, et al. (2000). "The Significance of the Classroom Effect in Primary Schools: An Application of Creemers' Comprehensive Model of Educational Effectiveness." *School Effectiveness and School Improvement* 11(4): 501-529.
- Landsheere, G. d. (1994). *Le pilotage des systemes d' education*. Bruxelles, De Boeck - Université.
- Lauder, H., I. Jamieson, et al. (1998). Models of effective school: Limits and capabilities. *School Effectiveness for whom?* R. Slee, G. Weiner and S. Tomlinson. London, Falmer Press.
- Law 2986 of the Greek Government (2002). Law 2986. FEK 24/13-02-02, *Journal of the government of the Greek democracy*: 231-242.
- Leney, T. (1999). European Approaches to Social Exclusion: Hodgson, A. Analysing Education and Training Policies for Tacking Social Exclusion, Kogan Page, London: 33-45
- Leo-Ruth (1991). Who decides? Policy makers in English Language Arts Education. *Handbook of Research on Teaching the English Language Arts*. New-York, Macmillan.
- Levine, D. U. and L. W. Lezotte (1990). *Unusually Effective Schools: A review and Analysis of Research and Practice*. Madison, WI, National Centre for Effective Schools Research and Development.
- Levine, D. U. (1990). Update on Effective Schools: Findings and Implementations from Research and Practice. *The Journal of Negro Education*, Autumn 1990, 59(4): 577-584

- Levine, D. U. (1992). An interpretative review of US research and practice dealing with unusually effective schools. *School Effectiveness: Research, policy and practice*. D. Reynolds and P. Cuttance. London, Cassell: 25-47.
- Levine, D. U. and L. W. Lezotte (1995). Effective schools research. *Handbook of research on multicultural education*. J. A. Banks and C. A. McGee. London, Prentice Hall International.
- Lezotte, L. and B. Bancroft (1985). "School Improvement Based on Effective Schools Research: A Promising Approach for Economically Disadvantaged and Minority Students." *Journal of Negro Education* 54(3): 301-312.
- Lezotte, L. (1986). *School effectiveness: Reflections and future directions*. American Educational Research Association, San Francisco.
- Lightbown, P. M. (2000). "Classroom SLA Research and Second Language Teaching." Oxford University Press 2000 21(4): 431-462.
- Little, A. and R. Willey (1981). *Multi-Ethnic Education: The Way Forward*. London, Schools Council Pamphlet: 38.
- Lockheed, M. E. and E. Hanushek (1994). Concepts of Educational Efficiency and Effectiveness, *Human Resources Development and Operations Policy*: 17.
- Lodge, C. M. (1996). The Meanings of 'ineffective' and 'failing' schools. *School Effectiveness and School - Improvement*. London, University of London: 92.
- Losardo, A. and A. Notari-Syverson (2001). *Alternative approaches to Assessing Young Children*. Baltimore, London, Toronto, Sydney, Paul Brookes Publishing.
- Louis, K. S. (1998). "Effects of Teacher Quality of Work Life in Secondary Schools on Commitment and Sense of Efficacy." *School Effectiveness and School Improvement* 9(1): 1-27
- Luyten, H. (1994). "School Size Effects on Achievement in Secondary Education . Evidence from the Netherlands, Sweden and the U.S.A." *School Effectiveness and School Improvement* 5(1): 75-99.
- Luyten, H. (2003). "The Size of School Effects Compared to Teacher Effects: An Overview of the Research Literature." *School Effectiveness & School Improvement* 14(1): 31-51.
- Maas, C., and Hox, J. (2002). "Sample Sizes for Multilevel Modelling." Department of Methodology and Statistics, Faculty of Social Sciences, Utrecht University.
- MacBeath, J. (1999). *Improving Education Outcomes: an international perspective*. PDN Conference for School Leaders.
- Macbeath, J. and P. Mortimore (2001). *Improving School Effectiveness*. Buckingham-Philadelphia, Open University Press.

- MacKenzie, D. (1983). "Research on school improvement: An appraisal of some recent trends." *Educational Researcher* 12(April): 5-16.
- Madaus, G., T. Kellaghan, et al. (1979). "The sensitivity of measures of school effectiveness." *Harvard Educational Review* (49): 207-230.
- Mandeville, G. (1988). "School effectiveness indicators revisited: cross year stability." *Journal of Educational Measurement* (25): 349-366.
- Marmot, M., Fuhrer R., Ettner S., Marks N., Bumpass L., and Ryff C. (1998). "Contribution of Psychosocial Factors to Socioeconomic Differences in Health. " *The Milbank Quarterly* 76(3), <http://www.milbank.org/quarterly/763feat.html>.
- Marshall, C. and G. B. Rossman (1989). *Designing Qualitative Research*. London, Sage.
- Martinez, P. (2000). *Raising achievement; a guide to successful strategies*. London, FEDA.
- Mason, D. A. and R. Burns (1997). "Reassessing the Effects of Combination Classes." *Educational Research and Evaluation* 3(1): 1-53.
- Maughan, B., Pickles, A., et al. (1990). "Can Schools Change? I. Outcomes at Six London Secondary Schools." *School Effectiveness and School Improvement* 1(3): 188-210.
- Mavrogiorgos, G. (1985). "Educational reforms, technocracy and technical control. The case of the new textbooks (in Greek)." *Contemporary education*: 21-31.
- McBer, H. (2000). *Research into teacher effectiveness*. London.
- McCallion, P. (1998). *Achievement for All*. London, The Stationary Office.
- McCarthy, J. (1998). *The Accelerated Schools Project - Working with culturally diverse at-risk populations in urban areas*. The International Congress on School Effectiveness and school improvement '98, Manchester.
- McPherson, A. (1992). *Measuring Value-added in schools*. National Commission on Educational Briefing, No 1, February 1992, London.
- McLaughlin, B. (1992). *Myths and misconceptions about second language learning: what every teacher needs to unlearn*. Santa Cruz, University of California, National Center for Research on cultural diversity and Second Language Learning: 9.
- McLellan, K., J. West, et al. (2003). "Beware the Jabberwock, my son! The jaws that bite, the claws that catch!" *Measuring school district outputs by school district inputs*. American Educational Research Association, Chicago.
- Meighan, R. and I. Siraj-Blatchford. (1997). *A sociology of educating*. London, Cassell.
- Meijnen, G. W., N. W. Lagerweij, et al. (2003). "Instruction characteristics and cognitive achievement of young children in elementary schools." *School Effectiveness and School Improvement* 14(2): 159-187.

- Meuret, D. and S. Morlaix (2003). "Conditions of Success of a School's Self-Evaluation: Some Lessons of an European Experience." *School Effectiveness and School Improvement* 14(1): 53-71.
- Miles, M. and M. Huberman (1984). *Qualitative data analysis: a sourcebook of new methods*. London, Sage.
- Ministry of Education Direction of Primary Education (1996). Ministerial Decree (Y.A. Φ. 2/ 206 /Γ1/1141/ of 2/08/1996). Athens, Ministry of Education Direction of Primary Education,.
- Ministry of Education Direction of Primary Education (1998). Ministerial Decree Φ13.1/ 767/ Γ1/884/3-9-1998. Programmes of creative activities taking place in whole day schools. Athens, Ministry of Education Direction of Primary Education.
- Ministry of Education Direction of Primary Education (1999). Ministerial Circular Φ13.1/717/ Γ1/742/21-9-1999. Programmes of creative activities taking place in whole day schools. Athens, Ministry of Education Direction of Primary Education.
- Ministry of Education Direction of Primary Education (1999) Ministerial Decree (Y.A. Γ1/708/99 (also published in ΦΕΚ 1789, Α part/ 28/9/1999), Athens, Ministry of Education Direction of Primary Education.
- Ministry of Education Direction of Primary Education (2000). Ministerial Circular Φ13.1/897/ Γ1/694/6-9-2000. Programmes of creative activities taking place in whole day schools. Athens, Ministry of Education Direction of Primary Education,; 3.
- Mok, M. (1995). "Sample size requirements for 2-level designs in educational research" *Multi-level Modelling Newsletter*, 7(2): 11-15.
- Mortimore, P., P. Sammons, et al. (1988). *School Matters: The Junior Years*. London, Open books.
- Mortimore, P. (1991). "School Effectiveness Research: Which way at the Crossroads?" *School Effectiveness and School Improvement* 2(3): 213-229.
- Mortimore, P. (1993). "School Effectiveness and the Management of Effective Learning and Teaching." *School Effectiveness and School Improvement* 4(4): 290-310.
- Mortimore, P. Sammons, P. and Thomas, S. (1994). "School Effectiveness and value added measures." *Assessment in Education: Principles, Policy and Practice* 1: 315-332.
- Mortimore, P. and G. Whitty (1997). *Can School Improvement Overcome the Effects of Disadvantage?* London, Institute of Education.
- Mortimore, P. (1998). *The Road to Improvement: Reflections on School Effectiveness*. Lisse, The Netherlands, Abingdon, Swets & Zeitlinger.
- Mortimore, P. and G. Whitty (1999). *School Improvement: A Remedy for Social Exclusion? Tackling Disaffection & Social Exclusion*. A. Hayton. London, Kogan Page: 80-94.

- Moursund, D. G. (1996). *Increasing your expertise as a problem solver: Some roles of computers*. Oregon, Eugene.
- Muijs, D. and D. Reynolds (2000). "School effectiveness and Teacher effectiveness in mathematics: some preliminary findings from the evaluation of the mathematics enhancement programme (primary)." *School effectiveness and school improvement*. 11(3): 273-303.
- Muijs, E. D. and D. Reynolds (2001). *Effective Teaching: Evidence and Practice*. London, Paul Chapman Publishing.
- Muijtaba, T. and P. Sammons (2000). '*Contextualising school performance at Key Stage 1 and Key Stage 2 2000*'. London, Institute of education, University of London: 16.
- Myers, K. and H. Goldstein (1997). Failing schools and failing system. *1997 ASCD yearbook: Rethinking Educational Change with heart and mind*. A. Hargreaves. Alexandria, Association for Supervision and Curriculum Development.
- Nash, R. (2001). "Progress and school effectiveness: non-cognitive dispositions and within-class markets." *Journal of Educational Policy* 16(2): 89-102.
- Nichols, J. D. (1996). "Cooperative Learning: A Motivational Tool to Enhance Student Persistence, Self-Regulation, and Efforts to Please Teachers and Parents." *Educational Research and Evaluation* 2(3): 246-260.
- Νικολάου, Γ. και Κοριλάκη Π. (2005). Η ενίσχυση της προσπάθειας των αλλοδαπών και παλιννοστούντων μαθητών στο Δημοτικό Σχολείο. Εμπειρική έρευνα σχετικά με τη λειτουργία των Τάξεων Υποδοχής & των Φροντιστηριακών Τμημάτων.
- Nowakowski, J., M. A. Bunda, et al. (1985). *A Handbook of Educational Variables. A Guide to Evaluation*. Boston - Lancaster, Kluwer - Nijhoff Publishing.
- Nunes, T. and P. Bryant (1997). *Learning and Teaching Mathematics. An International Perspective*. London, Psychology Press.
- Nuttall, D., H. Goldstein, et al. (1989). "Differential School Effectiveness." *International Journal of Educational Research: Developments of School Effectiveness Research* 13(7): 769-776.
- Nuttall, D. (1990). *Differences in examination performance*. London, Research & Statistics Branch, ILEA.
- Office for Standards in Education (Great Britain) (1994). *Educational support for minority ethnic communities : a survey of educational provision funded under section 11 of the 1966 Local Government Act*. London, Ofsted.
- Office for Standards in Education (Great Britain) (2002). *Achievement of Black Caribbean pupils: Good practice in secondary schools*. London, Ofsted. Office for Standards in Education (Great Britain) (2002). London, HMSO: 42 p.

- Office for Standards in Education (Great Britain) (2003). *Inspecting schools : framework for inspecting schools*, Office for Standards in Education.
- Office for Standards in Education (OFSTED) (1995). *Homework in Primary and Secondary Schools*. London, HMSO Publications Centre.
- Office for Standards in Education (OFSTED) (1996). *Homework in Primary and Secondary schools*. London, HMSO.
- Office for Standards in Education (OFSTED) (1998). *School Evaluation Matters*. London, DfEE: 36.
- Office for Standards in Education (OFSTED) (1999a). *Inspecting Subjects and Aspects 11-18: Equal Opportunities*. London, Office for Standards in Education: 6.
- Office for Standards in Education (OFSTED) (1999b). *Schools requiring special measures and schools with serious weaknesses: The Annual Report of Her Majesty's Chief Inspector of Schools*, OFSTED.
- Office for Standards in Education (Great Britain) (1999). *Raising the attainment of minority ethnic pupils; school and LEA responses*, London, OFSTED
- Office for Standards in Education (OFSTED) (2000a). *Inspecting Schools: The framework*. London, Office for Standards in Education: 44.
- Office for Standards in Education (OFSTED) (2000b). *Handbook for Inspecting Primary and Nursery Schools*. London, Office for Standards in Education.
- Office for Standards in Education (OFSTED) (2000c). *Inspection of Local Education Authorities: Grade criteria for Inspection Judgements*. London, Office for Standards in Education: 29.
- Office for Standards in Education (OFSTED) (2000d). *Helping Schools to Carry out self-evaluation*. London, Office for Standards in Education: 6.
- Office for Standards in Education (OFSTED) (2001). *Action Planning for school improvement*. London, www.OFSTED.gov.uk.
- Office for Standards in Education (OFSTED) (Great Britain) (1994). *Primary Matters*. London.
- Olthof, A. L. and E. Lugthart (1992). *Effectiveness of Various Types of Secondary School Organization: A Contingency Approach*. *School Effectiveness and Improvement*. J. Bashi and Z. Sass. Jerusalem, The Magness Press: 344-369.
- Ontario Ministry of Education (2003). *Individual Education Plans: Standards for Development, Program Planning, and Implementation*, © Queen's Printer for Ontario.
- Opdenakker, M.-C. and J. V. Damme (2000). "The Importance of identifying levels in multi-level Analysis: An illustration of the effects of ignoring the top or intermediate levels in school effectiveness research." *School effectiveness and school improvement*. 11(1): 103-130.

- Opdenakker, M.-C. and J. V. Damme (2001). "Relationship between school composition and characteristics of school process and their effect on Mathematics Achievement." *British Educational Research Journal* Vol. 27(4): 407-432.
- Ortiz, A. A., C. Y. Wilkinson, et al. (1991). *Aim for the Best: Assessment and Intervention Model for the Bilingual Exceptional Student*. Austin, Texas, The University of Texas at Austin Department of Special Education, College of Education.
- Osborne (2001). Unravelling Underachievement among African American boys from an identification with academics perspective. *Journal of Negro Education*, Howard University, Fall 2001, 98(4): 555-565.
- Owen, D. (1987). *Mathematics in Primary Education*. London, The Falmer Press.
- Palmaffy, T. (1998). "The Gold Star State." Policy review: 30-38.
- Paschal, R. A., Weinstein, T. & Walberg, H. (1984). The effect of homework on learning: a quantitative synthesis. *Journal of Educational Research* (78): 97-104.
- Patton, M. Q. (1987). *How to Use Qualitative Methods in Evaluation*. London, Sage.
- Patton, M. Q. (2002). *Qualitative Research and Evaluation Methods*. London, Sage.
- Pawson, R. and N. Tilley (1997). *Realistic Evaluation*. London, Sage.
- Pearse, J. (1986). *Standards and the LEA*. Windsor, Berkshire, NFER-NELSON.
- Perkins, D. (1995). *Outsmarting IQ: The emerging science of learnable intelligence*, The Free Press.
- Pettit G., Bates J., Dodge K, and Meece D. (1999) The Impact of After-School Peer Contact on Early Adolescent Externalizing Problems Is Moderated by Parental Monitoring, Perceived Neighborhood Safety, and Prior Adjustment. *Child Development*, 70(3), May-June 1999
- Pinnell, G. S., D. E. DeFord, et al. (1988). *Reading Recovery: Early Intervention for At-Risk First Graders*. Arlington, Virginia, Educational Research Service.
- Pinnell, G. S. (2000). *A complete copy of Reading Recovery: An Analysis of a Research-Based Reading Intervention*. Columbus, OH: Reading Recovery Council of North America.
- Plewis, I. (1991). "Underachievement: a case of conceptual confusion." *British Educational Research Journal* 17(4): 377-393.
- Plewis, I. (1991). "Pupils' progress in reading and mathematics during primary school: associations with ethnic group and sex." *Educational Research* 33: 133-140.
- Plewis, I. (1997). *Statistics in Education*. London, Arnold.

- Plewis, I. and J. Hurry (1998). "A Multilevel Perspective on the Design and Analysis of Intervention Studies." *Educational Research and Evaluation* 4(1): 13-26.
- Pollack, S., J. Chrispeels, et al. (1988). *A description of district factors that assist in the development of equity schools*. American Educational Research Association, New Orleans.
- Prandy K. (2002) *CAMSIS: Social Interaction and Stratification Scale*. Cardiff School of Social Sciences <http://www.cardiff.ac.uk/socsi/CAMSIS>
- Qualifications and Curriculum Authority (1997). *The value added National Project*. London, Qualifications and Curriculum Authority, QCA: 14.
- Rasbash, J., W. Browne, et al. (2000). *A user's guide to MLwiN manual (Second Edition)*. London, Institute of Education.
- Rassool, N. and L. Morley (2000). "School Effectiveness and the Displacement of Equity Discourses in Education." *Race, Ethnicity and Education* 3(3): 237-255.
- Raudenbush, S., B. Rowan, et al. (1991). "A multilevel, Multivariate model for studying school climate with estimation via the EM algorithm and Application to U.S. High-School Data." *Journal of Educational Statistics* 16(4): 295-330.
- Raudenbush, S. W. and D. Willms (1995). "The Estimation of School Effects." *Journal of Educational and Behavioural Statistics* 20(4): 307-335.
- Reynolds, D., B. Creemers, et al. (1989). *School Effectiveness and School Improvement. Proceedings of the first International Congress*, London. Cardiff, School of Education, University of Wales, University of Cardiff.
- Reynolds, D. (1991). "School effectiveness and school improvement in the 1990's." *Association for child's psychology and Psychiatry Newsletter* 13(2): 5-9.
- Reynolds, D. (1992). School effectiveness and School Improvement: An Updated Review of the British Literature. *School Effectiveness: Research, Policy and Practice*. D. Reynolds and P. Cuttance. London, Cassell: 1-24.
- Reynolds, D. and P. Cuttance (1992). *School Effectiveness: Research, Policy and Practice*. London, Cassell.
- Reynolds, D. (1993). "Linking School Effectiveness and School Improvement Practice: Towards a Synergy." *School Effectiveness and School Improvement* 4(1): 37-58.
- Reynolds, D. (1994). School Effectiveness and Quality in Education. *Improving Education: Promoting Quality in Schools*. P. Ribbins and E. Burrridge. London, Cassell: 18-31.
- Reynolds, D., B. P. M. Creemers, et al. (1994). *Advances in School Effectiveness Research and Practice*. Oxford, Pergamon.
- Reynolds, D., B. Creemers, et al. (1994). School Effectiveness: the need for an international perspective. *Advances in School Effectiveness Research and Practice*. D. Reynolds, B. Creemers, S. Stringfield and C. Teddie. Oxford, Pergamon: 183-201.

- Reynolds, D. (1996). Turning Around Ineffective Schools: Some Evidence and Some Speculations. *Merging Traditions: The Future of Research on School Effectiveness and School Improvement*. J. Gray, D. Reynolds, C. Fitz-Gibbon and D. Jesson. London, Cassell.
- Reynolds, D. and S. Farrell (1996). *Worlds apart? A review of International Surveys of Educational Achievement involving England*. London, OFSTED (Office for Standards in Education).
- Reynolds, D., P. Sammons, et al. (1996). "School Effectiveness and School Improvement in the United Kingdom." *School Effectiveness and School Improvement* 7(2): 133-158.
- Reynolds, D. and L. Stoll (1996). *Merging school effectiveness and school improvement: the knowledge bases. Making good schools: Linking school effectiveness and school improvement*. D. Reynolds, R. Bollen, B. Creemers et al. London, Routledge: 94-112.
- Reynolds, D., B. Creemers, et al. (1998). Climbing the Educational Mountain: Conducting the International School Effectiveness Research Project (ISERP). *Doing Research about Education*. G. Walford. London, Falmer Press: 111-124.
- Reynolds, D., C. Teddlie, et al. (2000). An Introduction to School Effectiveness Research. *The International Handbook of School Effectiveness Research*. C. Teddlie and D. Reynolds. London and New York, London and New York: 3-25.
- Reynolds, D. (2000). School Effectiveness: The International Dimension. *The International Handbook of School Effectiveness Research*. C. Teddlie and D. Reynolds. London and New York, Falmer Press: 232-256.
- Reynolds, D. (2000). "Should pedagogical change be mandated? Can and should pedagogical change be mandated at times?" *Journal of Educational Change* 1(2): 193-198.
- Reynolds, D., C. Teddlie, et al. (2000a). Linking School Effectiveness and School Improvement. *The International Handbook of School Effectiveness Research*. C. Teddlie and D. Reynolds. London and New York, Falmer Press: 206-231.
- Reynolds, D. and C. Teddlie (2000b). The future agenda for school effectiveness research. *The International Handbook of School Effectiveness Research*. London, Falmer Press: 322-343.
- Reynolds, D. (2001). Beyond School Effectiveness and School Improvement? *School effectiveness and school improvement: alternative perspectives*. A. Harris and N. Bennett. London, Continuum: 26-46.
- Reynolds, D., S. Stringfield, et al. (2002). The intellectual and policy context. *World class schools: international perspectives on school effectiveness* D. Reynolds, B. Creemers, S. Stringfield, C. Teddlie and G. Schaffer. London, Routledge / Falmer: 3-14.
- Reynolds, D., B. Creemers, et al. (2002). Creating World - Class Schools: what have we learned? *World class schools: international perspectives on school effectiveness*. D.

- Reynolds, B. Creemers, S. Stringfield, C. Teddlie and G. Schaffer. London, Routledge / Falmer: 2766-293.
- Riddel, A. R. (1988). *School Effectiveness in Secondary Education in Zimbabwe: a multilevel analysis*. Institute of Education. London, University of London: 322.
- Riddell, S. and S. Brown (1991). *School Effectiveness research: Its messages for school Improvement*. Edinburgh, HMSO.
- Robertson, P. and D. Toal (2001). Extending the quality framework: Lessons from case study schools. *Improving school effectiveness*. J. Macbeath and P. Mortimore. Buckingham - Philadelphia, Open University Press.
- Robson, C. (2002). *Real world research : a resource for social scientists and practitioner-researchers*. Oxford, Blackwell.
- Rosenfeld, L. B., Richman, J. M., and Bowen G. L. (1998). "Low social support among at-risk adolescents" *Social Work in Education* 20(4): 245-260.
- Rottman, R. B. D. (1995). *Class stratification*. Hemel Hempstead, Hertfordshire, Harvester Wheatsheaf.
- Rowe, K. J. (2000). "Assessment, League Tables and School Effectiveness: Consider the Issues and 'Let's Get Real!'" *Journal of Educational Enquiry* 1(1): 73-98.
- Rudd, P. (2001). *School Improvement through ICT: The Case of England*. The CIDREE Seminar: The Use of ICT in Supporting Effective Learning and Teaching, Scottish Education and Teaching Technology Show, SECC, Glasgow.
- Rudduck and Gray (1999). '*Gender and Achievement*'.
http://www.standards.dfes.gov.uk/genderandachievement/more_index.html, DfEE.
- Rudner, L. and W. Schafer (2001). *Reliability*, ERIC Digest. 2003.
- Ruth, L. (1991). Who decides? Policy makers in English Language Arts Education. *Handbook of Research on Teaching the English Language Arts*. J. Flood, J. M. Jensen, D. Lapp and J. R. Squire. New-York, Macmillan: 85-107.
- Rutter, M., B. Maughan, et al. (1979). *Fifteen Thousand Hours*. London, Paul Chapman Publishing.
- Rutter, M. (1983). "School Effects on Pupils' Progress: Research Findings and Policy Implications." *Child Development* 54(1): 1-29.
- Ruxton, S. (1996). *Children in Europe*. London, NCH Action for Children.
- Sack, J. L. (2000). "Education Department issues rules on using title I funds to improve schools." *Education week*: 3.

- Sammons, P., P. Cuttance, et al. (1993). "Differential school effectiveness: results from a re-analysis of the inner London Education Authority's School Project Data." *British Educational Research Journal* 19(4): 381-405.
- Sammons, P. (1994). Findings from School Effectiveness Research: Some Implications for Improving the Quality of Schools. *Improving education: Promoting Quality in Schools*. Ribbins, P. and Burrridge, E. London, Cassell.
- Sammons, P., S. Thomas, et al. (1994). *Assessing school effectiveness - a summary of the full report*. London, International School Effectiveness and School Improvement Centre, Institute of Education, University of London: 9.
- Sammons, P., S. Thomas, et al. (1994). *Assessing School Effectiveness: Developing Measures to put School Performance in Context*. Annual Conference of the BERA, Oxford, International School Effectiveness and Improvement Centre (ISEIC), Institute of Education, London.
- Sammons, P. (1995). "Gender, Ethnic and Socio-economic Differences in Attainment and Progress: a longitudinal analysis of student achievement over 9 years." *British Educational Research Journal* 21(4): 465-485.
- Sammons, P., S. Thomas, et al. (1995a). *Accounting for variations in academic effectiveness between schools and departments*. European Conference on Educational Research (ECER) and British Educational Research Association (BERA), Bath.
- Sammons, P., J. Hillman, et al. (1995b). *Key characteristics of Effective Schools: A review of school effectiveness research*. London, Office for Standards in Education and Institute of Education, University of London.
- Sammons, P., S. Thomas, et al. (1995c). *Understanding School and Departmental Differences in Academic Effectiveness*. International Congress of School Effectiveness and Improvement, Leewarden, the Netherlands.
- Sammons, P. (1996). "Complexities in the Judgement of School Effectiveness." *Educational Research and Evaluation* 2(2): 113-149.
- Sammons, P., P. Mortimore, et al. (1996). Do schools perform consistency across outcomes and areas? *Merging traditions: The future of research on school effectiveness and school improvement*. J. Gray, D. Reynolds, C. Fitz-Gibbon and D. Jesson. London, Cassell: 3-29.
- Sammons, P., A. West, et al. (1997). "Accounting for variations in pupil attainment at the end of Key Stage 1." *British Educational Research Journal* 23(4): 489-511.
- Sammons, P. and D. Reynolds (1997). "A partisan evaluation - John Elliot on School Effectiveness." *Cambridge Journal of Education* 27: 123-135.
- Sammons, P., R. Smees, et al. (1998). *The impact of background factors on pupil attainment and progress in Scottish schools*. International Congress for School Effectiveness & Improvement, Manchester.

- Sastry N., Ghosh-Dastidar, B., Adams J., and Pebley A. (2003). "The Design of a Multilevel Survey of Children, Families and Communities: The Los Angeles Family and Neighbourhood Survey." *Paper No. 2003-06, RAND Corporation, Santa Monica*, <http://opr.princeton.edu/papers/opr0306.pdf>: 30.
- Schagen, I. (1994). "Multilevel Analysis of the Key Stage 1 National Curriculum Assessment Data in 1991 and 1992." *Oxford Review of Education* 20(2): 163-171.
- Scheerens, J., Vermeulen, C. (1989). "Generalisability of Instructional and School Effectiveness Indicators Across Nations." *International Journal of Educational Research: Developments of School Effectiveness Research* 13(7): 789-799.
- Scheerens, J. and B. Creemers (1989). "Conceptualising School Effectiveness." *The International Journal of Educational Research* 13(7): 691-706.
- Scheerens J., and Creemers B. (1999). Review and prospects of educational effectiveness research in the Netherlands, in Roel J. Bosker, Bert P.M. Creemers and Sam Stringfield, 'Enhancing Educational Excellence, Equity and Opportunity' ICO, Kluwer Academic Publishers: 197-222
- Scheerens, J. (1992). *Effective Schooling: Research, Theory and Practice*. London, Cassell.
- Scheerens, J. (1993). "Basic School Effectiveness Research: Items for a research Agenda." *School Effectiveness and School Improvement* 4(1): 17-36.
- Scheerens, J. and B. Creemers (1996). "School Effectiveness in the Netherlands: The Modest Influence of a Research Programme." *School Effectiveness and School Improvement* 7(2): 181-195.
- Scheerens, J. and R. Bosker (1997). *The Foundations of Educational Effectiveness*. Oxford, Pergamon.
- Scheerens, J. (1997). "Conceptual Models and Theory Embedded Principles of Schooling." *School Effectiveness and School Improvement* 8(3): 269-310.
- Scheerens, J. and Creemers, B. (1999) 'Review and prospects of educational effectiveness research in the Netherlands' in Bosker, R., Creemers B., and Stringfield, S. 'Enhancing Educational Excellence, Equity and Opportunity' ICO, Kluwer Academic Publishers: 197-222
- Scheerens, J. Vermeulen, C. J., and Pelgrum, W. J. (1989). "Generalizability of instructional and school effectiveness indicators across nations." *International Journal of Educational Research* 13(7): 789-799.
- Scherer, M. (2001). "How and why standards can improve student achievement." *Educational leadership* 59(1): 14-18.
- Scheurich, J. J., L. Skrla, et al. (2000). "Thinking Carefully about Equity and Accountability." *Phi Delta Kappan* 82(4).

- Schmidt, W. H., C. C. McKnight, et al. (1997). *Many Visions, Many Aims: (TIMSS Volume 1): A Cross-National Investigation of Curricular Intentions in School Mathematics*. Dordrecht, Kluwer Academic Publishers.
- Schwandt, T. A. (1997). *Dictionary of Qualitative Inquiry*. London, Sage.
- Secada, W. G. (1992). Race, ethnicity, social class, language, and achievement in mathematics. *Handbook of Research on Mathematics Teaching and Learning*. D. A. Grouws. New York, Macmillan. 1: 623-660.
- Sensenbaugh, R. (1995). *Reading Recovery*, ERIC Digests. 2002.
- Sharp, C., W. Keys, et al. (2001). *Homework: A review of recent research*. The Mere, NFER.
- Sharp, S. and L. Croxford (2003). "Literacy in the first year of schooling: A multilevel analysis." *School Effectiveness and School Improvement* 14(2): 213-231.
- Shoemaker, J. (1984). *Research-based school improvement practices*. Hartford, Connecticut State Department of Education.
- Sidoli, A. (1992). *Friendship and Peer Tutoring: Paliative Prescription or Panacea?* Institute of Education. London, University of London: 153.
- Sijde, P. C. v. d. and W. Tomic (1992). "The influence of a teacher training programme on student perception of classroom climate." *Journal of Education for teaching* 18 (3): 287-295.
- Silverman, D. (2001). *Interpreting Qualitative Data*. London, Sage.
- Slavin, R. and N. Madden (1987). *Effective classroom programmes for pupils placed at risk'*. Baltimore, Centre for research on elementary and middle schools, Johns Hopkins University.
- Slavin, R. and M. Leighton (1990). *Success for All: Effects on the achievement of Limited English Proficient Children*. Baltimore, MD, Center for research on Effective Schooling for Disadvantaged Students, Baltimore, MD.
- Slavin, R. (1995). *Co-operative learning: Theory, research and practice*. Boston, London, Allyn and Bacon.
- Slavin, R., N. Madden, et al. (1996). "Success for All." *Journal of education for students placed at risk* 1(1): 41-76.
- Slavin, R. and N. Madden (1998). *Success for All / Exito Para Todos: Effects on the reading Achievement of students acquiring English*. Baltimore MD, Center for research on the education of students placed at risk, Office of educational research and improvement (ED).
- Slavin, R. E. and O. S. Fashola (1998). *Show me the evidence! : Proven and promising programs for America's schools*. London, Thousand Oaks.

- Slee, R. and G. Weiner (1998). *School Effectiveness for Whom? Challenges to the School Effectiveness and the School Improvement Movements*. London, Falmer Press.
- Smith, D. and S. Tomlinson (1989). *The school effect: A study of Multi-Racial comprehensives*. London, Policy Studies Institute.
- Smyth, E. (1999). "Pupil performance, Absenteeism and school drop-out: A multi-dimensional analysis." *School effectiveness and school improvement* 10(4): 480-502.
- Snijders, T. and R. Bosker (1999). *Multilevel Analysis*. London, Sage.
- Spours, K. and A. Hodgson (1996). *The formative value-added system*. London, Institute of Education.
- Steedman, J. (1983). *Examination results in Mixed and Single Sex Schools: Findings from the National Child Development Study*. Manchester; Equal Educational Opportunities Commission.
- Σταμέλος Γ. (2002), *Το ελληνικό εκπαιδευτικό σύστημα. Πρώτη και δεύτερη βαθμίδα. Δομές και ποσοτικά δεδομένα*, KEE, Αθήνα.
- Sternberg, R. (1988). *The triarchic mind: A new theory of human intelligence*, Penguin Books.
- St. John, N. (1970). 'The validity of children's reports of their parents' educational level: a methodological note', *Sociology of Education*, 43, Summer: 255-269.
- St. John, N. (1970). 'The validity of children's reports of their parents' educational level: a methodological note', *Sociology of Education*, 43, Summer: 255-269.
- Stanton D. (2003). Developing social protection for an enlarged European Union. *Social Statistics in the Enlarged EU: 89th DGINS Conference*, Athens, European Communities, Eurostat. Luxemburg: 43-51.
- Stattin H. and Magnusson D. (1995). Onset of official delinquency: its co-occurrence in time with educational, behavioural, and interpersonal problems. *British Journal of Criminology*: 35(3): 417-425.
- Stoll, L., D. Reynolds, et al. (1996). Merging school effectiveness and school improvement: Practical examples. *Making good schools: Linking school effectiveness and school improvement*. D. Reynolds, R. Bollen, B. Creemers et al. London, Routledge: 113-147.
- Stoll, L. and P. Mortimore (1997). School Effectiveness and School Improvement. *Perspectives on School Effectiveness and School Improvement*. J. W. a. M. Barber. London, Bedford Way Papers, Institute of Education, University of London: 9-24.
- Stoll, L. (1999). "Realising our potential: Understanding and developing capacity for lasting improvement." *School effectiveness and school improvement* 10(4): 503-532.
- Stradling, R., L. Saunders, et al. (1991). *Differentiation in action: A whole school approach in raising attainment*. Slough, NFER - DES.

- Strand, S. (1997). "Pupil Progress during Key stage 1: a value added analysis of school effects." *British Educational Research Journal* 23(4): 471-487.
- Strand, S. (1998). "A 'value added' analysis of the 1996 primary school performance tables." *Educational Research* 40(2): 123-137.
- Strand, S. (1999). "Ethnic group, Sex and Economic Disadvantage: associations with pupils' educational progress from Baseline to the end of Key Stage 1." *British Educational Research Journal* 25(2): 179-202.
- Stringfield, S., C. Teddlie, et al. (1985). "Classroom interaction in effective and ineffective schools: Preliminary results from phase III of the Louisiana School Effectiveness study." *Journal of classroom interaction* 20: 31-37.
- Stringfield, S. and C. Teddlie (1987). *A Time to Summarise: Six Years and Three Phases of the Louisiana School Effectiveness Study*. The American Educational Research Association, Washington DC, The American Educational Research Association.
- Stringfield, S. and C. Teddlie (1989). The first three phases of the Louisiana School Effectiveness Study. *School effectiveness and school improvement*. B. Creemers, T. Peters and D. Reynolds. Amsterdam, Swets and Zeitlinger.
- Stringfield, S., C. Teddlie, et al. (1992). A five-Year Follow - Up of Schools in the Louisiana School Effectiveness Study. *School Effectiveness and Improvement*. J. Bashi and Z. Sass. Jerusalem: 381-414.
- Stringfield, S. (1994). Outlier studies of school effectiveness. *Advances in School Effectiveness Research*. D. Reynolds, B. Creemers, P. Nesselrodt et al. Oxford, Pergamon.
- Stringfield, S. (1995). "Attempts to enhance students' learning: A search for valid programs and highly reliable implementation techniques." *School Effectiveness and School Improvement* 6(1): 67-96.
- Stringfield, S. and R. Herman (1996). "Assessment of the State of School Effectiveness Research in the United States of America." *School Effectiveness and School Improvement* 7(2): 159-180.
- Stuart, N. (1994). Quality in Education. *Improving Education*. P. Ribbins and E. Burrige. London, Cassell: 8-17.
- Sukhnandan, L. (1999). *An investigation into gender differences in achievement. Phase 1: a review of recent research and LEA information on provision*. Slough, NFER.
- Swain, M. and S. Lapkin (1982). *Evaluating Bilingual Education: A Canadian Case Study*, Clevedon, Avon.
- Swann, L., FRS, et al. (1985). *Education for All: The report of the Committee of Inquiry into the Education of Children from Ethnic Minority Groups*. London, Department of Education and Science: 807.

- Swann, L., FRS, et al. (1985). *Education for all: a brief guide to the main issues of the report*. London, Department of Education and Science: 14.
- Swartz, S. L. and A. F. Klein (1996). *Reading Recovery: An Overview*.
- Tani C. R., Chavez E. L., and Deffenbacher J. L.. (2001). "Peer isolation and drug use among white non-Hispanic and Mexican American Adolescents" *Addiction*. 36(141): 127-140.
- Tashakkori, A. and C. Teddlie (1998). *Mixed Methodology: Combining Qualitative and Quantitative approaches*. London, Sage.
- Tatar, M. and G. Honenzyk (1996). "Immigrant and host pupils' expectations of teachers." *British Journal of Educational Psychology* 66: 289-299.
- Teddlie, C., P. C. Kirby, et al. (1989). "Effective versus ineffective schools. Observable differences in the classroom." *American Journal of Education*. 97(3): 221-236.
- Teddlie, C. and S. Stringfield (1989). *Classroom differences in effective and ineffective schools*. Annual meeting of the American Educational research Association, San Francisco.
- Teddlie, C. and S. Stringfield (1992). *Case Histories from a Longitudinal Study of School Effects*. The American Educational Research Association, San Francisco, The American Educational Research Association.
- Teddlie, C. (1993). *Schools make a difference: lessons learned from a 10-year study of school effects*. London, Teacher's College Press.
- Teddlie, C. and S. Stringfield (1993). *Schools make a difference*. New York, Teachers College Press.
- Teddlie, C., D. Reynolds, et al. (2000). The Methodology and Scientific Properties of School Effectiveness Research. *The International Handbook of School Effectiveness Research*. C. Teddlie and D. Reynolds. London and New York, Falmer Press: 55-133.
- Teddlie, C. and D. Reynolds (2000). *The International Handbook of School Effectiveness Research*. London, Falmer Press.
- Teddlie, C. and D. Reynolds (2001). "Countering the critics: responses to recent criticisms of school effectiveness research." *School Effectiveness & School Improvement* 12(1): 41-82.
- Teddlie, C., D. Reynolds, et al. (2002). The methodology strategy of ISERP. *World class schools: International perspectives on school effectiveness*. D. Reynolds, B. Creemers, S. Stringfield, C. Teddlie and G. Schaffer. London, Routledge / Falmer: 15-32.
- Tesch, R. (1990). *Qualitative Research: Analysis types and Software Tools*. London, The Falmer Press.
- The National Commission on Excellence in Education (1983). *A Nation at risk*.

- The National Literacy Trust (2003). *The Government response to boys' underachievement*, The National Literacy Trust.
- The Offices of Her Majesty's Chief Inspectors of Schools in England and Wales (1995). *School Accountability and Evaluation in New Zealand: comparisons with England and Wales*. Cardiff, Offices of Her Majesty's Chief Inspector: 31.
- The Swedish National Agency for Education (2001). *Preconditions and pupil achievement at municipal and school level – an internet application*. The 14th International Congress for school effectiveness and school improvement, Toronto.
- Thomas, S., Sammons P., and Mortimore, P. (1995). Stability and consistency in secondary schools' effects on students' GCSE outcomes over three years, *School effectiveness and School Improvement*, 8: 169-197.
- Thomas, S. (1995). Considering primary school effectiveness: an analysis of 1992 Key Stage 1 results, *The curriculum journal*, 6(3): 279-295.
- Thomas, S., Sammons, P., et al. (1995). *Accounting for variations in academic effectiveness between schools and departments*. ECER, Bath.
- Thomas, S. and P. Mortimore (1996). "Comparison of value-added models for secondary school effectiveness." *Research papers in Education* 11(1): 5-33.
- Thomas, S., P. Sammons, et al. (1997). "Differential Secondary School Effectiveness: comparing the performance of different pupil groups." *British Educational Research Journal* 23(4): 451-469.
- Thomas, W. P. and V. Collier (1997). "School Effectiveness for Language Minority Students." *NCBE Resource Collection Series* 9: 96.
- Thorntone, S. J. (1991). Teacher as curricular instructional safe-keeper in social studies. *Handbook of research on social studies, teaching and learning*. J. P. Sharer. New York, MacMillan.
- Thurpp, M. (1997). "How school mix shapes school processes: a comprehensive study of New Zealand's schools." *New Zealand Journal of Educational Studies* 33(2): 53-82.
- Thurpp, M. (1999). *Schools Making a Difference: Let's be realistic! School Mix, School Effectiveness and the social limits of reform*. Buckingham, UK, Open University Press.
- Thurpp, M. (2001). "Recent School Effectiveness Counter-Critiques: problems and possibilities." *British Educational Research Journal* 27(4): 443-457.
- Tizard, B., Blatchford, P., Burke, J., Farquhar, C., & Plewis, I. (1988) *Young Children at School in the Inner City*, Hove, Lawrence Erlbaum Associates.
- Tollefson, J. (1991). *Planning Language, Planning Inequality*, New York. 2002.
- Tomlinson, S. and H. Tomes (1983). *Ethnic minorities in British Schools : a review of the literature, 1960-82*. London, Aldershot : Gower.

- Torrance, H. (1997). Assessment, Accountability and Standards: Using Assessment to Control the Reform of Schooling. *Education: Culture, Economy, Society*. A. Halsey. Oxford: 320-331.
- Tsoukalas, K. (1977). *Dependence and reproduction. The social role of educational mechanism in Greece (1820-1922) (in Greek)*. Athens, Themelio.
- Tymms, P. (1992). "The Relative Effectiveness of Post 16 Institutions in England (Including Assisted Places Scheme Schools)." *British Educational Research Journal* 18(2): 175-192.
- Tymms, P. (1997). "Science in primary schools: an investigation into differences in the attainment and attitudes of pupils across schools." *Research in science and technological education* 15(2): 149-159.
- Tymms, P. Merrell, C. and Henderson, B. (1997). "The first year in school: a quantitative investigation of the attainment and progress of pupils." *Educational Research and Evaluation* 3: 149-159.
- Tymms, P. (1999). *Baseline assessment and monitoring in primary schools: achievements, attitudes and value-added indicators*. London, David Fulton.
- USA, C. House, et al. (2001). *No Child Left Behind Act*: conference report to accompany H. R. 1, Washington D.C, U.S. Government Printing Office 2001.
- U.S. Department of Health and Human Services (2003). *Strengthening Head Start: What the Evidence Shows: Report, Human Services Policy*, <http://aspe.hhs.gov/hsp/StrengthenHeadStart03/index.htm>, Washington D.C, U.S., June 2003.
- Valli, L., D. Cooper, et al. (1997). Professional Development Schools and Equity: A Critical Analysis of Rhetoric and Research. *Review of Research in Education*. M. W. Apple. Washington, DC, American Educational Research Association. 22: 251-270.
- Van De Jong M. (1989). Educational Climate and Achievement in Dutch Schools, in *School Effectiveness and Improvement: Proceedings of the first International Congress* Reynolds D., Creemers B., and Peters T. *School Effectiveness and Improvement: Proceedings of the first International Congress* London, RION.
- Veenman, S. (1997). "Combination Classes Revisited." *Educational Research and Evaluation* 3(3): 262-276.
- Velzen, W. V., M. Miles, et al. (1985). *Making school improvement work: a conceptual guide to practice*. Leuven, Belgium, Acco Publishers.
- Verdis, A. (2002). *School Effectiveness and Educational Evaluation in Greece*. Institute of Education. London, University of London: 388.

- Walberg, H. (1984). "Improving the productivity of American Schools." *Educational Leadership* 41: 19-27.
- Walberg, H. (1986). Synthesis of research on teaching. *Handbook of research on teaching*. M. C. Wittrock. London, Macmillan: 214-229.
- Walberg, H. (1992). Measuring Educational Effectiveness. *School effectiveness and Improvement*. J. Bashi and Z. Sass. Jerusalem, The Magness Press.
- Walford, G. (1994). *Choice and Equity in education*. London, Cassell.
- Walsh, W. B. and N. E. Betz (1985). *Tests and Assessment*, Prentice Hall.
- Walsh, B. and N. Betz (1995). *Tests and Assessment*. New Jersey, Prentice Hall Inc.
- Webb, N. (1982). "Student Interaction and Learning in Small Groups." *Review of Educational Research* 52(3): 421-445.
- Werf, G. V. d. (1997). "Differences in School and Instruction Characteristics between High-, Average-, and Low-Effective schools." *School Effectiveness and School Improvement* 8(4): 430-448.
- West, M. and J. Beresford (1998). *Issues of Equity and diversity in the classrooms: Some of the problems confronting school improvement programmes*. International congress on school effectiveness and improvement, Manchester.
- Westerhof, K. J. (1992). "On the effectiveness of Teaching: Direct Versus Indirect Instruction." *School Effectiveness and School Improvement* 3(3): 204-215.
- White, J. and M. Barber (1997). *Perspectives on School Effectiveness and School Improvement*. London, Institute of Education, University of London.
- Whitty, G., S. Power, et al. (1998). *Devolution and choice in Education*. London, Open University Press.
- Whitty, G., Power, S., Gamarnikow, E., Aggleton P., Tyrer P. and Youdell D. (1999). Health, housing and Education: Tackling Multiple Disadvantage: Hodgson, A. Analysing Education and Training Policies for Tackling Social Exclusion, Kogan Page, London: 67-79
- Whitty, G. (2001). "Education, social class and social exclusion." *Education Policy*, 16(4): 285-295.
- Wiersma, W. and S. G. Jurs (1985). *Educational Measurement and Testing*. Boston, London, Sydney, Toronto, Allyn and Bacon.
- Willard Hartup (1992). Having Friends, Making Friends, and Keeping Friends: Relationships as Educational Contexts, ERIC Digest, ED345854, ERIC Clearinghouse on Elementary

and Early Childhood Education Urbana IL,
http://www.ericfacility.net/databases/ERIC_Digests/ed345854.html.

- Willms, D. (1986). *Do Scottish Education Authorities Differ in their Examination Results?* Edinburgh, The Scottish Council for Research in Education.
- Willms, D. (1992). *Monitoring school performance*. London, The Falmer Press.
- Willms, D. J., J. Douglas, et al. (2001). "Family, classroom and school effects on children's educational outcomes in Latin America." *School Effectiveness and School Improvement* 12(4): 409-445.
- Woods, D. and M. Cribb (2001). *The role of LEAs in school improvement. Effective LEAs and school improvement: making a difference*. D. Woods and M. Cribb. London, Routledge / Falmer: 1-14.
- Yang, M., H. Goldstein, et al. (1999). "The Use of Assessment Data for School Improvement Purposes." *Oxford Review of Education* 25(4): 469-483.
- Young, D. and B. Fraser (1993). "Socioeconomic and Gender Effects on Science Achievement: An Australian Perspective." *School Effectiveness and School Improvement* 4(4): 265-289.
- Young, D. J. (1998). *Characteristics of Effective Rural Schools: A Longitudinal Study of Western Australian Rural High School Students*. AERA (American Educational Research Association), San Diego, California.

